DEVELOPMENT AND TESTING OF BIOMONITORING TOOLS FOR STREAM MACROINVERTEBRATES IN THE LAKE TAHOE BASIN

Leska S. Fore
Statistical Design
136 NW 40th St.
Seattle, WA 98107
leska@seanet.com

Final Report

February 2007

Prepared for:

USDA-Forest Service
Lake Tahoe Basin Management Unit
Forest Plan Revision Team
35 College Drive
South Lake Tahoe, CA 96150

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Stream macroinvertebrates collected from tributary streams in the Lake Tahoe Basin were evaluated for their association with patterns of human land use. During the late summer of 2003, benthic macroinvertebrates were systematically collected from 171 locations on 10 tributary streams in the Tahoe Basin. Biological metrics related to taxa richness and relative abundance of key groups were tested against two independent measures of human disturbance. Metrics were more consistently correlated with measures of disturbance made at the site than with measures of development derived from remotely sensed data.

Of the 17 candidate metrics tested, 7 were strongly associated with site disturbance: number of stonefly, caddisfly, long-lived, intolerant, clinger, and EPT taxa and the percent of taxa that were non-insects. An additional index (O/E) used by the state of California to compare the number of observed vs. expected taxa was also correlated with disturbance measures. Metrics selected for the Tahoe Basin were similar to those identified as reliable indicators in other regional studies and metric values indicated that streams in Tahoe Basin were in better condition than in other areas of California.

Metrics were transformed to unitless scores ranging from 0-10 and scores were added to obtain an overall index, the Tahoe Basin multimetric index (MMI). Because human development was more intense near the lake shore, elevation and human disturbance were confounded. For this reason, metrics were tested for their association with human disturbance and elevation simultaneously using multiple regression. Most metrics and both indexes were significantly associated with both. The MMI was more highly correlated with disturbance than its component metrics or the O/E index.

The MMI developed in this study can be used to define regional expectations (reference condition) for Tahoe Basin tributary streams. A basin-wide probabilistic survey could be implemented during the 2007 field season that would yield information on status of stream condition and similar sampling over time could be used to monitor for trend. During the 2007 field season, reference sites in low elevation and low gradient habitat should be sampled and replicate samples should be collected to estimate MMI variance. A formal sampling plan should be designed that balances the needs for status, trend, and targeted monitoring.

ACKNOWLEDGMENTS

Shane Romsos (USFS) provided data sets, background materials, and administrative support. Bruce Davidson (USFS) provided maps, land use information, and additional measures of human influence. Sean Parks (USFS-PSW) calculated the development index from land use information. Travis Gallagher, Jennifer Slaughter, and Camisha Booth collected field samples and habitat data. Bob Wisseman (Aquatic Biology Associates) identified the macroinvertebrates. Barak Shemai (USFS) provided administrative support. Andy Rehn (CA DFG) provided CalEDAS information. Ian Chan (Garcia and Associates) assisted in directing field data collection efforts and Becky Parker (Garcia and Associates) calculated biological metrics. Chuck Hawkins (Utah State) provided O/E scores. Discussions with Shane Romsos and James Karr greatly improved this document.

INTRODUCTION

Numerous agencies in the Tahoe Basin are mandated to manage and protect the natural resources of the basin. All these agencies share a need to understand how management actions and activities, environmental policy, and patterns of residential and commercial development affect Lake Tahoe's ecosystem. To develop that understanding, scientific methods and protocols for monitoring and assessment are needed. An appropriate approach is specifically needed to understand the condition of streams and their associated riparian corridors.

Within the regulatory context of the Clean Water Act, states are required to assess their surface waters and report their condition to Congress (Davies and Jackson, 2006; EPA, 2006a). Over the years, a shift from primarily chemical monitoring to a more integrative assessment of the biological condition of surface waters reflects the understanding that the resident biological community is the best indicator of ecosystem health (Karr, 1991; Karr, 2006a). To support state monitoring and assessment, EPA has developed numerous guidance documents and specifically recommends monitoring of biological assemblages because (Barbour et al., 1999):

- Biological communities integrate the impact of different types of human activities.
- Communities integrate human influence through time.
- Biological monitoring is typically less inexpensive than chemical or toxicity testing.
- The status of biological communities is of direct interest to the public as a measure of a pollution free environment.

For streams, there are typically three choices for the biological assemblage to sample: vertebrates (fish, amphibians, and reptiles), macroinvertebrates (primarily insects), and algae (typically diatoms). Biological monitoring based on fish sampling is problematic in the western states due to low numbers of species and restrictions on sampling due to species listings as threatened or endangered. Compared with macroinvertebrates, algae are more directly associated with in-stream condition (e.g., nutrient levels) than with measures of habitat or riparian condition (Fore, 2003). In contrast, macroinvertebrates are most commonly sampled and every state has protocols for macroinvertebrate bioassessment (EPA, 2002).

Macroinvertebrate communities, composed largely of aquatic insects, live in streams year-round and variations in community composition and abundance have been shown to integrate the effects of human disturbances that cause changes in water chemistry, hydrologic flow patterns, habitat condition, and riparian vegetation (Kerans and Karr, 1994; Karr, 1998; Karr and Chu, 1999; Hodkinson and Jackson, 2006). Many macroinvertebrate taxa are year-round residents of streams; consequently, they integrate all the human influences occurring over space (upstream and adjacent to the stream corridor) and through time. Macroinvertebrates are

abundant in streams, simple to collect, relatively inexpensive to identify, and much is known about their specific life history traits and tolerances (Merritt and Cummins, 1996; Barbour et al., 1999). The unique tolerances of different species have been related to different types of human disturbance and have the potential to distinguish between different causes of degradation (Simon, 2002). They are flexible indicators in that they respond to a variety of different types of disturbance, are cost effective, and can be applied in a variety of contexts including site restoration, effectiveness of best management practices (BMP's), and evaluation of point sources (Karr and Chu, 1999; Vowell, 2001; Morley and Karr, 2002).

Like macroinvertebrates, many fish, amphibian, mammal, bird, and plant populations associated with stream riparian areas depend on the quality and quantity of water to carry out essential life history functions. Thus, biological monitoring of stream macroinvertebrates can protect these species as well. Macroinvertebrates directly support other populations as a primary food source for both aquatic and terrestrial vertebrates. Nakano and Murakami (2001) found that for streams in northern Japan, aquatic prey accounted for 26% of the annual energy demand for the entire bird assemblage. The presence of American dippers, a semi-aquatic bird, in Wyoming was associated with the types of macroinvertebrates typical of healthy streams (Feck and Hall, 2004). Riparian lizards allowed access to stream riparian areas grew faster due to the inclusion of aquatic insects in their diet (Sabo and Power, 2002).

Designing a monitoring plan

The primary purpose of monitoring and assessment of natural resources is to determine whether their condition has improved, declined, or remains unchanged. Ward et al. (1986) describe five steps that support a successful monitoring program. The focus of this classic paper was on chemical monitoring, but their five steps can easily be adapted to biological monitoring. An effective monitoring program must: 1) identify information expectations by defining program goals and objectives, 2) develop statistical design criteria, 3) develop a sampling plan, 4) develop written protocols for data handling, analysis, and management, and 5) report results of monitoring.

Under Step 1 (*Identify information expectations*), the Forest Service, in collaboration with other basin agencies and in response to regional planning efforts (TIIMS, 2007), will provide the monitoring framework and specific monitoring goals and objectives for assessing stream and riparian condition under the auspices of the Biological Resources Monitoring and Evaluation Program (BRMEP). The goal of this program is to identify specific monitoring goals and resource condition targets.

Step 2 (*Develop statistical design criteria*) translates the goals and hypotheses of Step 1 into an overall survey design. Elements of this step include identifying the population (e.g.,

stream sites within the Tahoe Basin), selecting a statistical survey method (e.g., probabilistic sampling from a linear stream model), and identifying specific statistical models and tests for evaluating the data (e.g., linear regression for trend detection).

Step 3 (*Develop a sampling plan*) is even more specific and identifies exactly what type of data will be collected in the field and where sample sites will be located. For Tahoe Basin streams, macroinvertebrates will be collected from each site. Other ancillary data related to habitat condition or human disturbance may also be collected and would be finalized as part of this step. Here a specific sampling plan is derived from the more general survey design developed in Step 2 and the exact location of sampling sites are defined along with a schedule for visiting them. Sites may be visited every year or in a recurring pattern with different panels of sites visited in successive years.

Under Step 4 (*Develop written protocols*), all procedures must be standardized and protocols documented so that data can be collected and analyzed consistently through time even if the personnel change. Much of this step is complete for Tahoe Basin streams because field protocols for macroinvertebrate collection have been defined and tested by others (Carter and Resh, 2001; Frazier et al., 2005; Herbst and Silldorff, 2006; Rehn et al., 2007). Data analysis methods used to derive an index of biological integrity from macroinvertebrate sample data are described in this document (see below). What remains to be accomplished is the implementation of data management methods to archive data and automate analysis procedures.

Step 5 (*Report results*) completes the monitoring cycle and much of this step should be completed before any data are collected. Information reporting procedures should specify a format for the monitoring information (e.g., written reports or interactive web pages), a distributional list for the report, and a review process to determine whether the data and analysis meet the information needs outlined in Step 1. In short, have the data answered the questions? The Forest Service and Tahoe Regional Planning Authority (TRPA) are required to produce an evaluation document every five years that summarizes the status of the indicators relative to established standards or targets. Macroinvertebrate sampling and reporting for stream condition should be matched to this evaluation schedule.

Report structure

This document primarily addresses elements under Steps 2, 3, and 4. Pilot data collected from Tahoe Basin tributary streams in 2003 was used to evaluate macroinvertebrate response to human disturbance in Tahoe Basin. The following sections (Methods, Results, and Discussion) describe the development and testing of biological indicators derived from macroinvertebrate samples. The final section (Recommendations) addresses the larger context and makes specific recommendations for implementing a long-term monitoring program in Tahoe Basin for

assessing the status and trends of the biological condition of streams. The recommendations are presented in a framework that follows the steps outlined above.

METHODS

This section summarizes the methods used to collect, process, and analyze the land use and biological data. In brief:

- Macroinvertebrates were collected from 171 locations on 10 streams in the field;
- Macroinvertebrate samples were electronically subsampled in the lab to standardize sample sizes;
- Human disturbance was measured in the field at the site scale and remotely at various spatial scales;
- Metrics were tested for correlation with measures of human disturbance;
- Metrics were tested simultaneously for association with both human disturbance measures and elevation using multiple regression.

Study sites

Benthic macroinvertebrates were collected from 10 streams within the Lake Tahoe Basin (Figure 1). Samples were collected systematically, starting from a random riffle near the stream mouth, moving upstream. The frequency or interval of sampling depended on the number of riffles that occurred on a given stream, with a maximum target of 30 samples per stream. Thus, riffles selected for sampling varied depending on the total length of stream and number of documented riffles on a given stream. The number of sampling locations along each stream ranged from 4 to 27 for a total of 171 (Table 1). On average, sampling sites were about 0.5 km apart. Sampling sites ranged in elevation from lake level at approximately 6200 ft to 8450 ft above sea level. Many samples were collected in stream sections that were generally steep with percent slope > 10% (76 out of 171 sites). Stream slopes along Tallac, Taylor, Upper Truckee, and the lower reaches of Trout (~10 sites) had a lower gradient and the reaches here were more meandering. Sites were visited in the late summer and early fall of 2003.

Table 1. Stream name, average percent slope, slope range for all sample sites, minimum and maximum elevation for all sample sites, average distance between sites, number of sites sampled on each stream (and number of sites with additional information to evaluate site condition).

Stream name	Average % slope	Min-max elev. (ft)	Distance b/n sites (km)	Number of samples
Blackwood	10.2 (2-43%)	6250-7350	0.4	27 (9)
Edgewood	16.2 (3-29%)	6250-7000	0.3	16 (6)
General	14.7 (2-44%)	6250-7000	0.9	9 (4)
Meeks	14.7 (2-32%)	6250-7000	0.5	13 (5)
Tallac	8.2 (3-12%)	6250-6450	0.5	8 (3)
Taylor	5.1 (3-11%)	6250-6400	0.3	11 (4)
Third	14.3 (3-56%)	6250-8450	0.3	31 (10)
Trout	13.6 (0-31%)	6250-8300	0.8	27 (9)
Upper Truckee	3 (2-5%)	6200-6250	0.6	4 (0)
Ward	7.5 (2-18%)	6250-6850	0.3	25 (9)
All sites	**11.7**	**6200-8450**	**0.5**	**171**

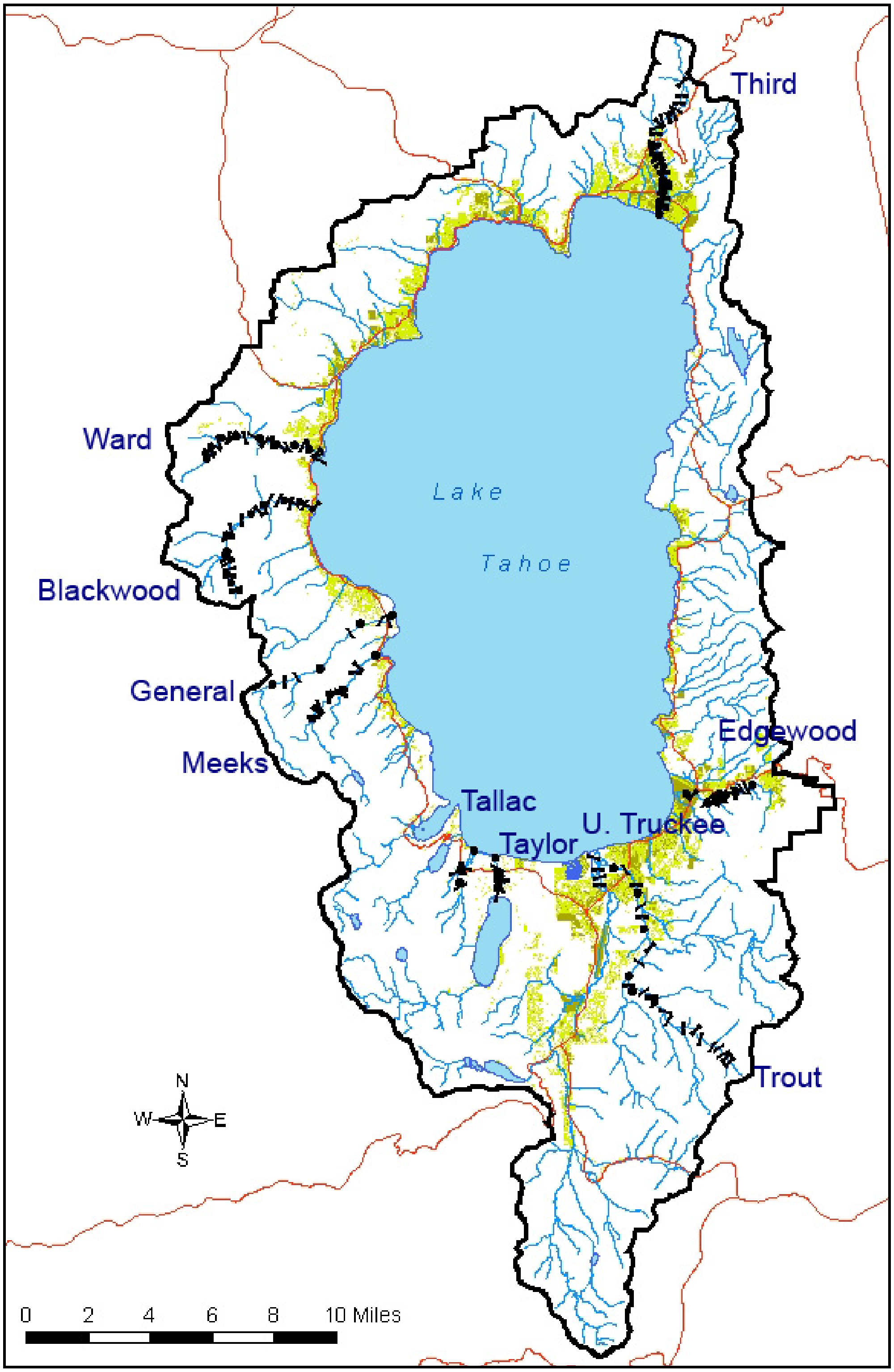

Figure 1. Lake Tahoe tributary sampling locations. Black lines indicate macroinvertebrate sampling only; black circles indicate macroinvertebrate, habitat, vegetation, and water quality sampling. Intensity of human development indicated by a gradient of green, where darker shading indicates more intense development.

Quantifying human disturbance

Most streams were characterized by more intense human development in the downstream areas closest to Lake Tahoe. Human disturbance was quantified in two different ways. The first method used satellite data to calculate the percentage of pixels that indicated human development within an area upstream of the sampling location (n = 171). The area upstream used to quantify percent development upstream ranged from a 30 x 50 m area to the entire upstream catchment. The second approach summarized human disturbance at the site scale based on eight observations made at the macroinvertebrate collection sites. Only a subset of the 171 sites had this information (n = 59).

Landscape scale

Three digital data sources were used to summarize the intensity of human development within stream watersheds: a parcel-based land use layer, transportation layers, and digital aerial photographs. Any area that was permanently cleared of natural vegetation, e.g., areas with buildings, roads, or golf courses were defined as "developed." Because downhill ski runs retain their native vegetation, although in an early seral condition, they were excluded. A 30 x 30 m pixel size was used to summarize development.

A parcel map created by the TRPA was used to define 90 types of land use, e.g., campgrounds, single family dwelling, multi-family dwelling, or airfield (Table 2). The average proportion of developed area for each parcel type was calculated from aerial photos for a subset of parcels and that percentage was applied to similar land use types recorded for other areas in the basin. As an example, a parcel identified as single family dwelling had on average, 51% of its area developed. For every 30 x 30 m pixel labeled as single family dwelling, 51 of the 100 possible 3-m cells were randomly selected and labeled as developed. Thus, each 30-m pixel was assigned a value from 0-100, which was an estimate of the proportion of the pixel that was developed.

The percent developed area upstream of each sampling site was estimated at seven different spatial scales: the 30 m pixel within which the sampling site was located, a buffer area 50 m upstream and 15 m on each side of the stream, a rectangular area 100 x 200 m oriented upstream, circles with radii of 200, 500, and 1000 m with the most downstream point on the sampling location and the area of the circle oriented along the stream channel, and for the entire watershed upstream (Figure 2). The percent developed area at each of these spatial scales was used to summarize human land use upstream of each sampling location.

Table 2. Selected examples of the 90 land use types and their associated percentage of developed land area.

Land use type	Average % Developed
Summer home	14
Developed campgrounds	22
Transmission and receiving facilities	36
Outdoor recreation concessions	48
Single family dwelling	51
Public health and safety facilities	65
Multiple family dwelling	77
Golf courses	78
Vehicle and freight terminals	89
Small scale manufacturing	90
Airfield	100

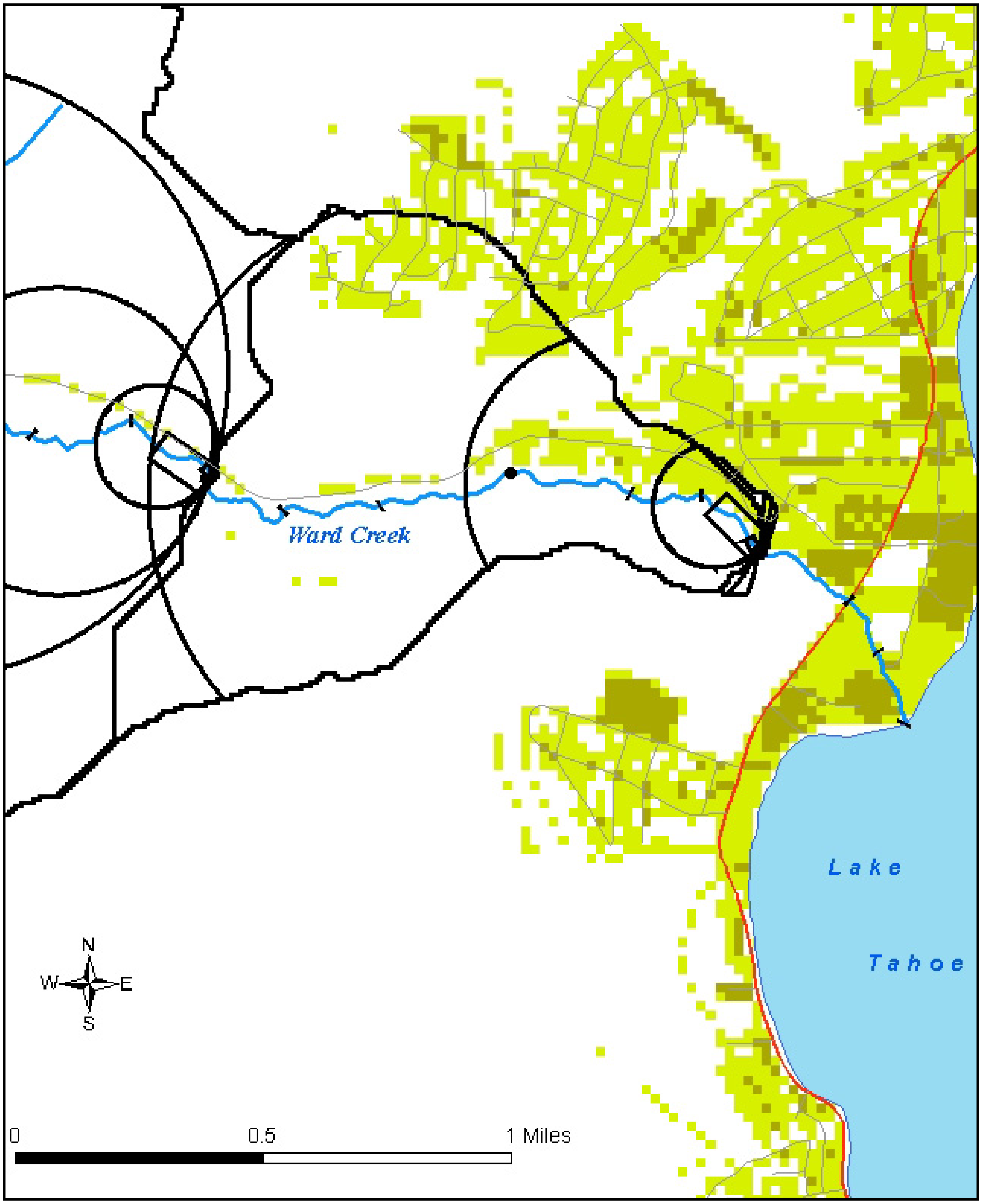

Figure 2. Detail from Ward Creek showing the intersection of the watershed and the areas upstream of the sampling site used to estimate percent development (darker green shading indicates more development; see Table 2). Shown are the 30 x 50 m rectangle, 100 x 200 m rectangle, and 200, 500, and 1000 m radius circles clipped to the watershed boundary.

Site scale

For a subset of sampling locations (59 out of 171), additional information was collected. Within each stream, information on site condition was collected at approximately every third or fourth site. Dissolved oxygen, pH, temperature, and conductivity were recorded along with habitat information summarizing the physical condition of the reach (Frazier et al., 2005). Habitat measures included presence of channel modification, a measure of percent erosion along the banks 15 m upstream of the sample site, whether large woody debris was present, whether human structures could be seen from the sample site, and percent canopy cover.

To summarize the intensity of human disturbance at the site scale, eight measures of site condition were used (Table 3). The eight site scale measures were coded as 0, 1, or 2 depending on whether the value observed for the site indicated low, moderate, or high disturbance. The scores were summed to provide an overall measure of site disturbance. The values for site disturbance could potentially range from 0 to 11 but for this data set ranged from 0 (least disturbed) to 7 (most disturbed).

Table 3. Eight measures recorded at the sample sites used to quantify site disturbance. Shown are the scoring rules used to combine the measures into a single overall index of site disturbance: 0 = low disturbance, 1 = moderate disturbance, and 2 = high disturbance. Scores were summed to obtain an overall measure of site disturbance.

	Low	**Moderate**	**High**
Site condition measure	**0**	**1**	**2**
Number of road crossings within 50 m upstream	0	1	
Percent development at the site (30 x 30 m pixel from satellite data)	< 30%	30-60%	> 60%
Conductivity	< 80	>= 80	
Channel modification	no	yes	
Percent erosion	< 30%	30-60%	> 60%
Large woody debris	yes	no	
Human structures present	no	yes	
Percent canopy	> 60%	30-60%	< 30%

Macroinvertebrate collection and identification

Macroinvertebrates were collected at a randomly selected longitudinal segment within a riffle habitat unit (roughly represented as either upper, middle, or lower third of the unit). Three collection locations were selected to represent the microhabitat types observed perpendicularly across the riffle (generally 2 margin samples and a thalweg sample). A D-frame kick net with a 500-um mesh net was placed on the substrate and an area of approximately 2 ft^2 was disturbed to a depth or 4–6 in for 2 minutes. Contents of the three kick net samples were combined. A minimum of 500 macroinvertebrates were identified by a professional taxonomist to the lowest practical taxonomic level, typically species for most insects including chironomids.

Although the target sample size was 500 individuals, many samples had very high numbers of individuals identified. Because taxa richness increases as a function of the number of individuals identified from a sample (Larsen and Herlihy, 1998), a random selection routine was used to electronically subsample the data from each site to 500 individuals (for metrics) or 300 individuals (for predictive models; Walsh, 1997). Macroinvertebrates were identified to lowest practical taxonomic level; for insects this was typically genus or species (including chironomids).

Data analysis

Taxa testing

Macroinvertebrate taxa for California have been identified as tolerant or intolerant based on best professional judgment of regional experts (CDFG, 2006). For this study, 57 out of 59 sites with site disturbance information were used. Two sites very close to the lake were excluded due to the potential inclusion of lentic species. Although tolerance and intolerance to general disturbance typically varies at the genus level, these data were too sparse to test genera. Instead taxa were grouped by family for taxa testing. The 57 sites were first ranked according to site disturbance, and then by percent development measured at a 50 x 30 m scale. Taxa were evaluated graphically for a decrease or increase with disturbance.

Metric and index calculation

Biological metrics are measures of biological condition that respond predictably and reliably to independent measures of human disturbance. For this study, metrics were evaluated that have been previously shown to reliably respond to many types and aspects of human disturbance in northern California and the Pacific northwest (Table 4). Multimetric indexes (MMIs) are calculated by converting metric values to unitless scores and summing the scores to obtain an overall assessment of biological condition (Karr and Chu, 1999). The benthic index of

biological integrity (B-IBI) and the northern California MMI are composed of metrics selected on the basis of their correlation with multiple independent measures of human disturbance (Karr, 1998; Karr and Chu, 1999; Rehn et al., 2005a; Rehn et al., 2007). The B-IBI developed for the Pacific Northwest and Japan includes 10 metrics selected for their association with urbanization, timber harvest, recreation, and pollutants (Appendix A; Fore et al., 1996; Karr, 1998; Karr and Chu, 1999; Fore et al., 2001; Booth et al., 2004). The natural history information used to calculate the metrics was derived from the California Department of Fish and Game's bioassessment database (CDFG, 2006).

Table 4. Metrics used in the Pacific northwest B-IBI and the northern California multimetric index, their description, and their documented response to human disturbance.

Metric	Description	Demonstrated response	PNW B-IBI	CA MMI
# Total taxa	Total number of unique taxa	Decrease	Yes	
# Ephemeroptera	Number of mayfly taxa	Decrease	Yes	
# Plecoptera	Number of stonefly taxa	Decrease	Yes	
# Trichoptera	Number of caddisfly taxa	Decrease	Yes	
# Long-lived	Number of taxa with life spans >1 year	Decrease	Yes	
# Intolerant	Number of intolerant taxa (tolerance values = 0–3)	Decrease	Yes	
% Tolerant organisms	Percent of organisms belonging to tolerant taxa (tolerance values = 7–10)	Increase	Yes	
% Predators	Percent of organisms that feed on other organisms	Decrease	Yes	Yes
# Clinger	Number of clinger taxa	Decrease	Yes	
% Dominance	Percent of organisms in the three most abundant taxa	Increase	Yes	
# EPT Taxa	Number of mayfly, stonefly, and caddisfly taxa	Decrease		Yes
# Coleoptera	Number of beetle taxa	Decrease		Yes
# Diptera	Number of true fly taxa	Decrease		Yes
% Intolerant organisms	Percent of organisms belonging to intolerant taxa	Decrease		Yes
% Non-Gastropod Scrapers	Percent of organisms that graze periphyton (excluding snails and limpets)	Decrease		Yes
% Shredder taxa	Percent of taxa that shred organic matter	Decrease		Yes
% Non-insect taxa	Percent of taxa that are not insects	Increase		Yes

The seven metrics included in the northern California MMI were selected from a larger set of 77 candidate metrics based on their correlation with watershed disturbance, agricultural cover in the watershed, road density, percent sand and fines, conductivity, phosphorus, and channel alteration (Rehn et al., 2005a).

California state agencies use both multimetric indexes (MMI) and O/E index values for site assessment and reporting purposes. Both indexes were developed for data sets collected from a much larger area than the Tahoe Basin that included very few sites from Tahoe Basin. O/E index values derived from a predictive model developed for California were also tested here (Hawkins et al., 2000; Ode and Rehn, 2005; Herbst and Silldorff, 2006).

O/E values were derived from a multivariate model that predicts the expected taxa based on natural features that successfully distinguished between reference sites. The predictive model for California requires 300-count samples and Tahoe Basin samples were subsampled electronically to meet that target. For Lake Tahoe Basin, taxa were predicted on the basis of average yearly temperature and watershed area. The index value for each site (O/E) was calculated as the number of predicted taxa observed at the site (O) divided by the number of taxa expected to occur (E). O/E values were based on a probability of capture > 0.5 for reference site taxa.

Metric and index testing

Metrics, O/E, and the Tahoe Basin MMI were tested for correlation with site disturbance and percent development at multiple spatial scales. Several metrics that were significantly correlated with disturbance were also correlated with elevation; therefore, metrics and indexes were tested again using multiple regression.

Nonparametric correlation was used to test for an association between biological measures and measures of disturbance (Spearman's r). Because statistical significance of a correlation coefficient (r) is a function of the sample size, a more strict criteria for metric selection was used than simple statistical significance. For large data sets (e.g., $n = 100$) a correlation coefficient of 0.17 will be statistically significant (alpha = 0.05, 1-sided test). Such a small correlation coefficient may be statistically significant but biologically not very meaningful.

Criteria for metric correlation was an r-value > 0.4 (or < –0.4). Metrics were also evaluated graphically to ensure that sites with known disturbance all had lower metric values (or higher values if that pattern was associated with disturbance). In contrast, a wider range of values was tolerated for sites with minimal disturbance because every source of disturbance cannot be captured (Rehn et al., 2005b).

Although macroinvertebrate data were collected at 171 locations and satellite data were available for all sites, only 59 sampling locations had information to calculate site condition. Of these 59, 2 were excluded due to their close proximity to the lake (< 100 m). Another 6 sites were excluded from correlation testing because they were all located at a higher elevation (> 7200 ft) where almost no human disturbance occurred. Other elevations had a mix of disturbance levels and these 6 sites were excluded to avoid confounding the analysis with sites that were high elevation with no site disturbance.

Human development and disturbance was higher closer to the lake shore. In this way, human disturbance was confounded with elevation. To ensure that biological measures were reliably associated with site condition, a multiple regression model was used to simultaneously test their association with disturbance and elevation. Multiple regression is mathematically very similar to standard linear regression, the difference is that rather than a single predictor variable, one or more variables may included to predict the variable of interest. Multiple regression was used to evaluate metrics that were significantly correlated with human disturbance. For each metric, one multiple regression model was used to evaluate the predictive power of three independent variables (site disturbance, percent development, and elevation).

To combine metrics into a MMI for Lake Tahoe Basin, metric values were converted to unitless scores. Metrics were scored from 0–10 using the 10^{th} and 90^{th} percentile values to define the best metric score (10) and worst score (0). For values in between, a linear transformation was used. As an example, if the 10^{th} percentile of a metric were equal to 2 and the 90^{th} percentile equal to 10, those values would be scored as 0 and 10 respectively. Values in between would be matched such that a metric value of 6 would receive a score of 5, a metric value of 3 would be scored as 2.5, and so on. Data from all 171 stream samples were used to calculate percentiles used to score the metrics. Because the Lake Tahoe Basin MMI was the sum of the six scored metrics, the sum was divided by 0.6 to yield a final MMI that ranged from 0–100.

RESULTS

This section describes the results of testing taxa, metrics, and indexes against an independent measure of human disturbance. In brief:

- More human disturbance was found closer to the lake (at lower elevation);
- Metrics were more highly correlated with site disturbance variables collected at the sample site than percent development values remotely derived at any spatial scale;
- Some metrics were correlated with both site disturbance and elevation; and
- Multiple regression indicated that site disturbance was the best predictor of biological condition.

Quantifying human disturbance

Percent development measured at the four largest spatial scales (circles with radii of 200, 500, and 1000 m, and entire upstream watershed) were highly correlated with each other (Spearman's r = 0.68–0.96, n = 51). In contrast, agreement between the various measures of percent development and site condition was poor ($r < 0.5$). Correlation between site disturbance and percent development *increased* slightly when percent developed was measured at larger spatial scales, the opposite of what was expected. These results suggest that summary measures of percent disturbance were consistent across spatial scales, but that site condition at the smallest spatial scale could not be reliably estimated using satellite data. Although measures of percent development were available for all sampling locations, only the 51 sites with information available to calculate the site disturbance index as well were used for correlation analysis so that correlation coefficients could be compared for the different measures of disturbance.

Taxa testing

Six families showed a tendency to decline with disturbance (Figure 3) while one family increased. The California database (CDFG, 2006) reported tolerance values for genera ranging from 0 indicating most intolerant to 10 indicating most tolerant. Tolerance value for families were based on the average of the tolerance values for genera found in the Tahoe Basin streams. The four caddisfly (Trichoptera) families that declined (and their tolerance values) were Apataniidae (0.5), Glossosomatidae (0.3), Rhyacophilidae (1.5), and Uenoidae (1). Two other families also declined, one a stonefly (Plecoptera), and the other a true fly (Diptera): Perlidae (1.5) and Thaumaleidae (2). One additional family increased with disturbance: Pisidium (8), the fingernail clam. Additional statistical testing was not attempted because the data set was too small; that is, many families had a small number of individuals.

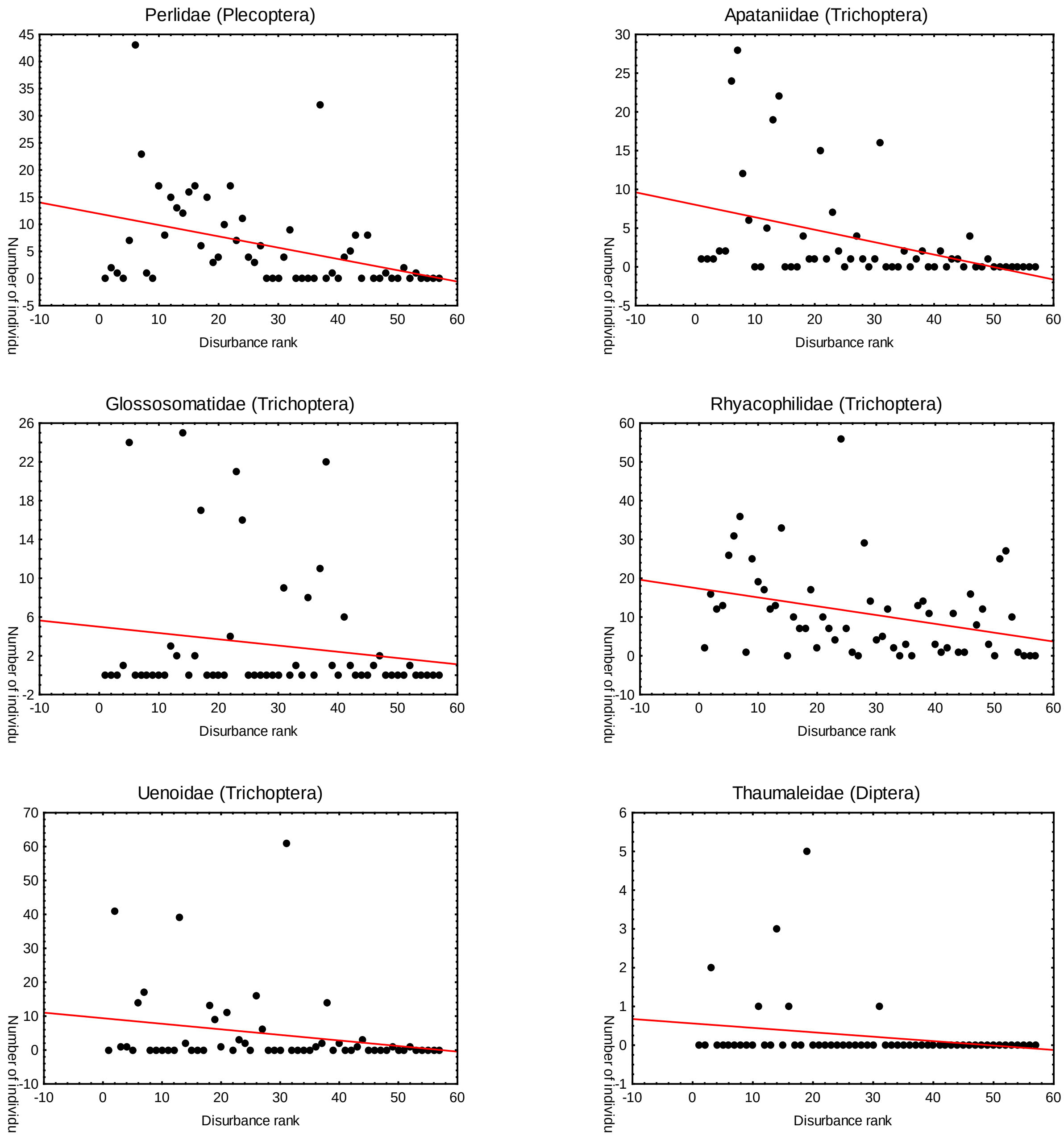

Figure 3. Six families that declined as site disturbance increased (n = 59 sites).

The influence of elevation

A primary concern with this study design was the potentially confounding relationship between elevation and human disturbance. Human development and activities tend to be closer to the lake shore where elevation is lowest and terrain is more flat. Percent developed area increased for sites closer to the lake shore, but site disturbance was not as closely associated with proximity to the lake. In general, greater site disturbance was observed at lower elevations; however, when elevation was divided into four categories, a range of site disturbance was observed in each category. An exception to this pattern was observed for the highest elevation sites which had the lowest site disturbance (Figure 4). When sites with elevation >7200 ft were excluded, correlation between elevation and site disturbance went from –0.42 to –0.26 (Spearman's *r*). For this reason, these six high elevation sites were excluded from additional analysis. The two most downstream sites on Blackwood and Ward were also excluded from subsequent analysis due to their close proximity to Lake Tahoe and the potential for different taxa associated with lentic habitats.

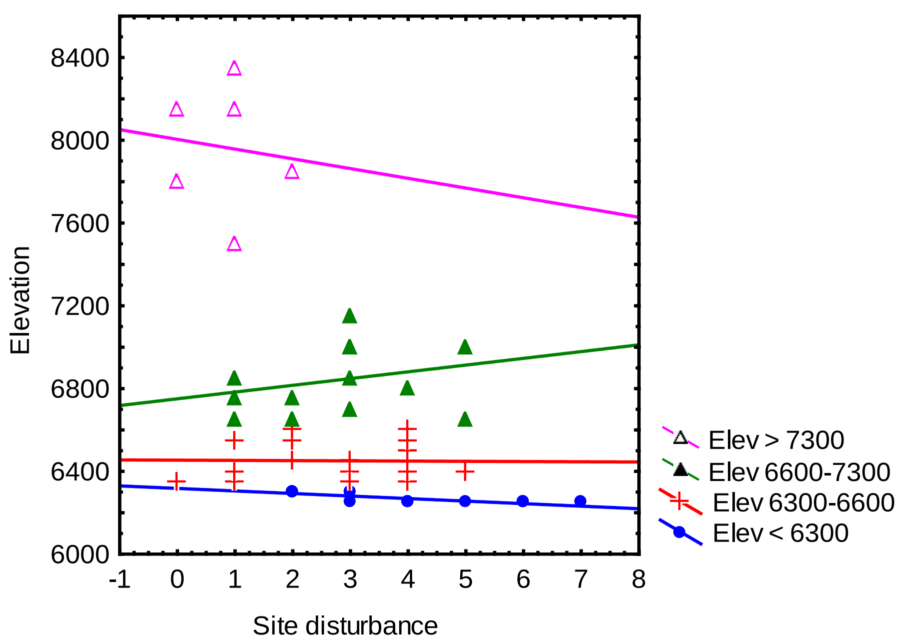

Figure 4. Relationship between site condition and elevation for four classes of elevation. Regression lines are shown separately for four ranges of elevation. When all sites were considered, elevation and site disturbance were correlated (Spearman's $r = –0.46$); however, correlation was due to a few high elevation with low site disturbance (see pink open triangles). Within the three lower ranges of elevation, a broader range of values for site disturbance was observed and elevation and site disturbance were not correlated for lower elevation sites.

Macroinvertebrate response to disturbance

Of the 17 metrics tested, 7 were highly correlated with the site disturbance index while 4 were correlated with percent development measured at the 500 m radius or watershed scale (Table 5). Of the 7 metrics that were correlated with site disturbance, 5 were B-IBI metrics and 2 were from the northern California MMI. Ten metrics were highly correlated with elevation. All the metrics that were correlated with disturbance were also correlated with elevation, with one exception (% non-insect taxa).

Of the 7 metrics that were highly correlated with site disturbance, 6 were taxa richness metrics: stonefly, caddisfly, long-lived, intolerant, clinger, and EPT taxa all declined as site disturbance increased (Figures 5–7). Metrics based on the percentage of individuals failed to show as strong a response to disturbance. Percentage of individuals that were tolerant, predators, non-gastropod scrapers, or non-insects failed to show a strong association with disturbance measured at any spatial scale. Percent of taxa that were non-insects was highly correlated with site disturbance and percent development at the watershed scale. The O/E index was highly correlated with elevation, site disturbance, and percent development measured at the 500 m radius and watershed scales.

Six of the seven metrics that were significantly correlated with site disturbance were converted to unitless scores and the sums of these scores defined the Tahoe Basin MMI for each sample (Table 6). Although significantly correlated with site disturbance, one of the seven metrics (EPT taxa richness) was not included because stonefly (Plecoptera) and caddisfly (Trichoptera) richness were included separately and EPT richness would represent a redundant measure of biological condition.

Table 5. Correlation between elevation, biological metrics and indexes, site disturbance, and percent development at three spatial scales (30 x 50 m, 500 m radius, and upstream watershed). Only *r*-values > 0.4 or < –0.4 are shown (Spearman's *r*, N = 51). All values shown were statistically significant ($p < 0.01$).

Metric or Index	Elevation	Site disturbance	%Dev 30 x 50 m	%Dev 500 m R	%Dev WS
Elevation	1.00				
# Total taxa	0.51			-0.42	
# Mayfly taxa	0.51				
# Stonefly taxa	0.60	-0.50		-0.50	-0.48
# Caddisfly taxa	0.46	-0.59		-0.41	
# Long-lived taxa	0.58	-0.45			
# Intolerant taxa	0.66	-0.58			
% Tolerant					
% Predator					
# Clinger taxa	0.41	-0.55			
% Dominance (3 taxa)					
# EPT taxa	0.67	-0.58		-0.46	-0.45
# Beetle taxa					
# True fly taxa					
% Intolerant					
% Non-gastropod scrapers	0.42				
% Shredder taxa	0.46				
% Non-insect taxa		0.44			0.42
O/E	0.43	-0.41		-0.42	-0.54
MMI	0.57	-0.65		-0.45	-0.42

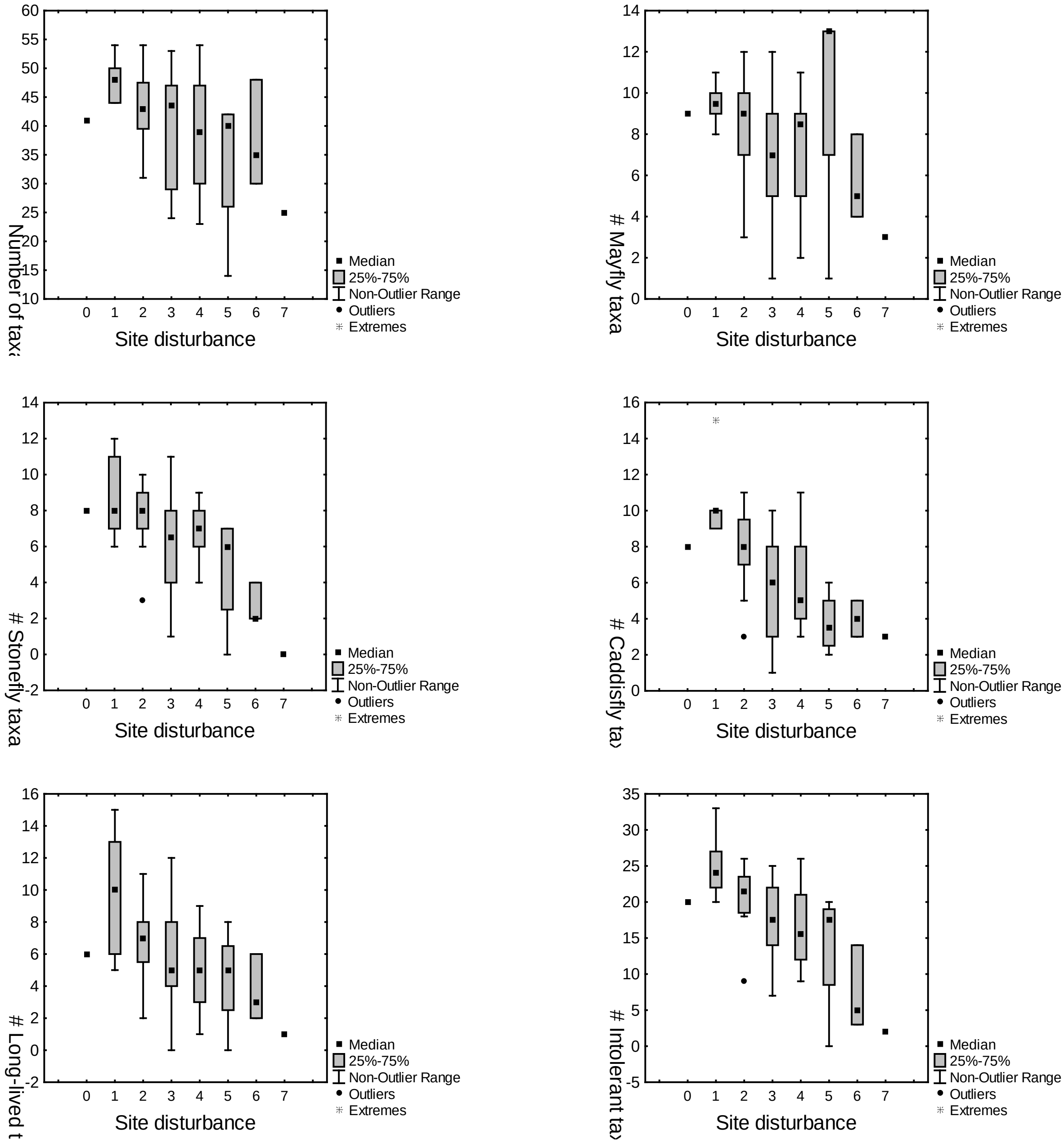

Figure 5. Five of these six taxa richness metrics declined consistently with site disturbance; mayfly taxa richness was an exception (N = 51 sites).

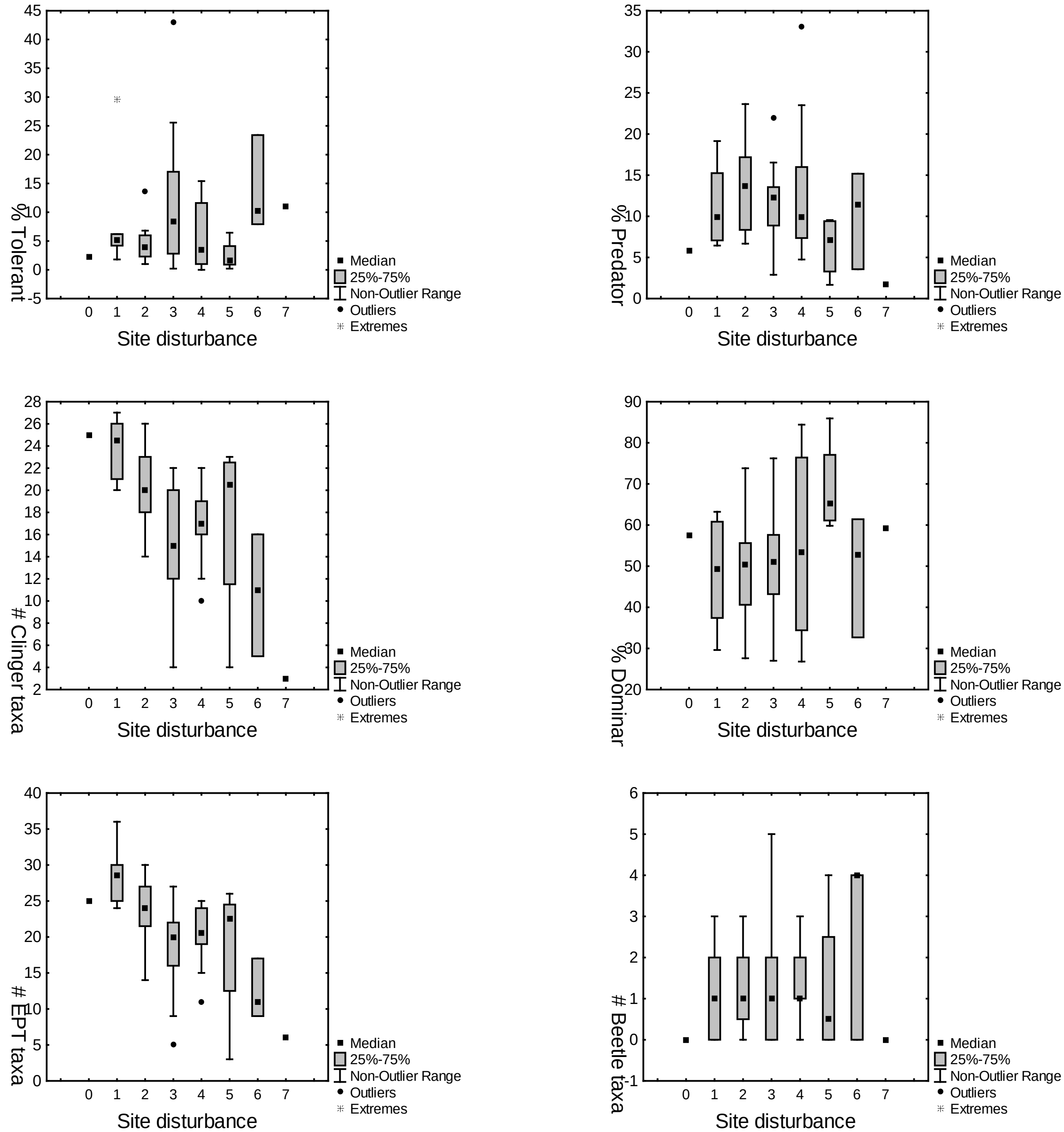

Figure 6. Percent tolerant individuals increased somewhat for the most disturbed sites only. Neither percent predators, percent dominance or beetle taxa were associated with site disturbance. Clinger and EPT taxa declined sharply with site disturbance (N = 51 sites).

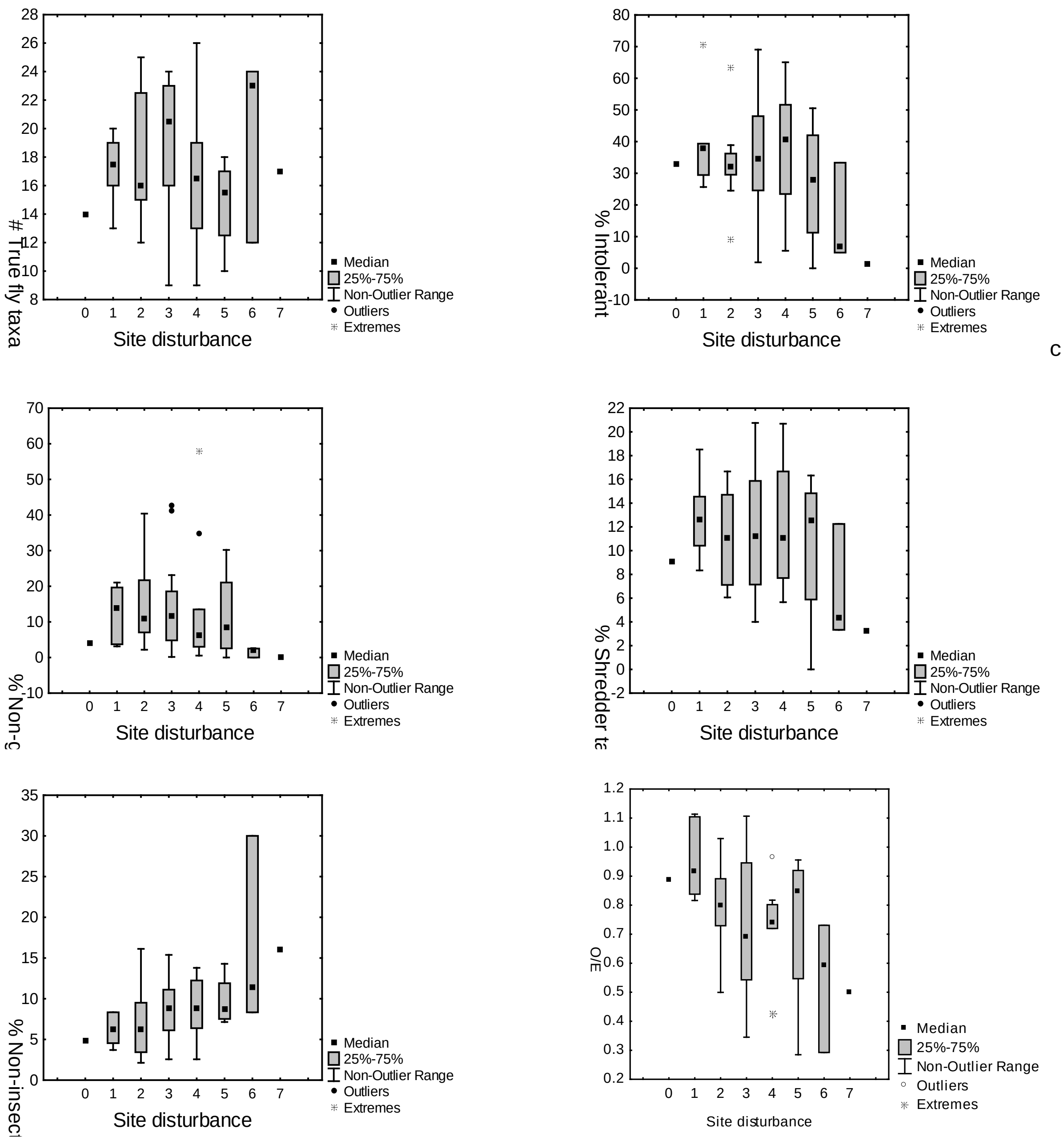

Figure 7. True fly taxa (Diptera), % non-gastropod scrapers, and percent shredder taxa failed to show a strong response to site disturbance. Percent intolerant individuals declined slightly. Percent non-insect taxa increased with disturbance and O/E declined as expected (N = 51).

Because several metrics were highly correlated with both elevation and site condition, multiple regression was used to test whether the biological metrics would continue to be significantly correlated with site condition when elevation was considered simultaneously. Multiple regression tested whether differences in biological metric and index values were better explained by changes in elevation or human disturbance. Nine separate multiple regression models were tested, one for each of the seven metrics that were significantly correlated with disturbance either at the site scale or subwatershed scale and one model each for O/E and the Tahoe Basin MMI (Table 7). Three predictor variables were included in each model: site disturbance, percent development (500 m radius), and elevation. All metrics and both indexes were significantly associated with site disturbance and elevation. An exception was percent non-insect taxa which was only associated with site disturbance and not elevation. Only one metric (caddisfly taxa) had a significant association with percent developed area as well.

Table 6. Rules for transforming six metrics to unitless scores for calculation of the Tahoe Basin MMI. For metric values <10th percentile, score = 0; for metric values >90th percentile, score = 10. Percent non-insect taxa was an exception and scoring was reversed (n = 171).

Metric	Mean	10th %tile	90th % tile	Scoring rule
# Stonefly taxa	6.9	3	10	(X – 3) * 10/7
# Caddisfly taxa	6.9	3	10	(X – 3) * 10/7
# Long-lived taxa	6.7	2	11	(X – 2) * 10/9
# Intolerant taxa	18.9	9	26	(X – 9) * 10/17
# Clinger taxa	18.4	11	25	(X – 11) * 10/14
% Non-insect taxa	8.3%	3.6	13.0	10 – ((X – 3.6) * 10/9.4)

Table 7. Beta values (standardized coefficients) from multiple regression for three predictors (site disturbance, elevation, and percent development for a 500 m circle), the adjusted R^2 for the model, and the *F*-value, n = 51). Statistical significance for the beta values is indicated as * $p < 0.01$, * * $p < 0.05$. All *F*-values for the full models were significant ($p < 0.01$).

	Total Taxa	Taxa Stone	Taxa Caddis	Taxa Long	Taxa Intol	Taxa Clinger	% Taxa Non-ins	O/E	MMI
Predictor variables									
Site disturbance	-0.29*	-0.45* *	-0.45* *	-0.43* *	-0.57* *	-0.56* *	0.49* *	-0.36* *	-0.56* *
Elevation	0.37* *	0.48* *	0.24*	0.48* *	0.49* *	0.24*		0.30*	0.38* *
%Dev 500m			-0.24*						
Full model									
Adjusted R^2	0.28	0.53	0.42	0.41	0.59	0.36	0.20	0.24	0.53
$F_{(3,55)}$	7.47	19.70	13.29	12.72	25.29	10.57	5.27	6.14	19.95

For multiple regression models, the adjusted R^2 represents how well the predictive variables, taken together, explained the variation in the dependent variable. For these nine models (one for each biological measure), the adjusted R^2 explains how much of the variance in MMI values was explained by site disturbance, elevation, and percent development. Although all nine models tested were statistically significant, the amount of variance they explained ranged from 20–59%. The best models (highest adjusted R^2) were observed for stonefly taxa, intolerant taxa, and the MMI. Lowest values were observed for percent non-insect taxa and O/E.

Tahoe Basin multimetric index

The multimetric index included six metrics selected for their correlation with site disturbance (see Table 6 above). The MMI was also highly correlated with the site disturbance index (r = –0.65) and somewhat correlated with percent development (r = –0.45 [500 m circle]; Figure 8A). The lower correlation between MMI and percent development was primarily due to high MMI values observed at sites downstream of golf courses (Figure 8B). Golf courses were assigned relatively high values for percent developed (78%, see Table 2). Sites downstream of golf courses had higher MMI values than expected based on their percent developed area upstream. Therefore, either MMI was not sensitive to the types of disturbances associated with golf courses or the percent developed values associated with golf courses (78%) were too high.

An alternative explanation for the high MMI values observed for golf courses relates to nutrient enrichment. High altitude stream systems, like those found in the Lake Tahoe Basin, are characterized by lower nutrients and more oligotrophic conditions. Golf courses typically add fertilizer to golf greens, which can run off into streams when it rains (Winter et al., 2002). Higher nutrients have been associated with an increase in mayfly taxa (Miltner and Rankin, 1998), and a similar pattern emerged for the Tahoe Basin. Sites with golf courses upstream had a much higher number of mayfly taxa present (Figure 9). All the sites with golf courses upstream were located along Third except one at the most downstream site on Edgewood. An increase in mayfly taxa due to disturbance is the opposite response expected to general disturbance and provides an example of identifying the stressor based on the biological assemblage (Simon, 2002).

The MMI was lower for sites <6400 ft, but did not increase for higher elevations (6400–8400 ft; Figure 8C). Because human influence was concentrated in these areas as well, this pattern of correlation was less of a concern than if, instead, MMI values had shown a tendency to increase at higher elevations where human disturbance was minimal.

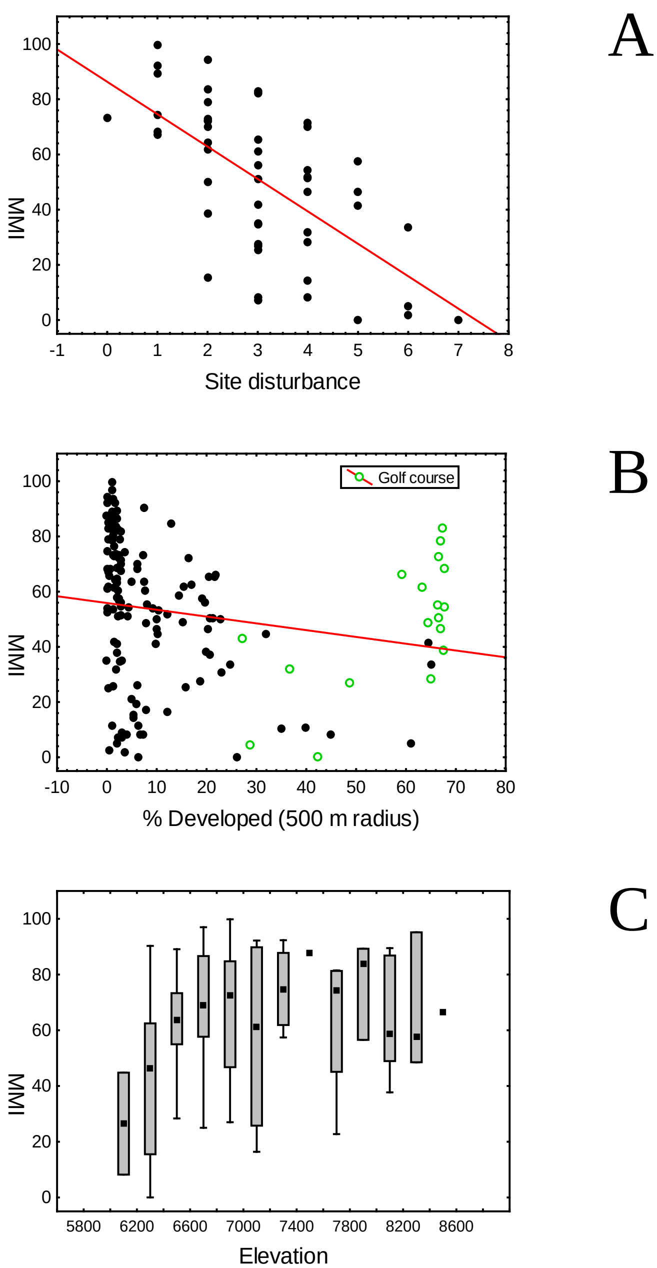

Figure 8. A) MMI was highly correlated with site disturbance (r^2 = 0.42, N = 51). B) MMI was not as highly correlated with percent development due to high MMI values for sites with golf courses located upstream (r^2 = 0.04, n = 150, elevation < 7200 ft). C) MMI was correlated with elevation at low elevation sites where most human development occurred.

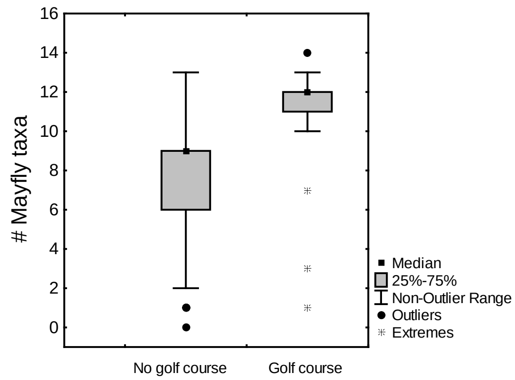

Figure 9. More mayfly taxa were found at sites with golf courses upstream (N=150, elevation < 7200).

The highest MMI values (indicating the best biological condition) were observed in Blackwood (reach 4), Meeks (reach 4), Trout (reaches 4, 5, and 7), and Ward (all reaches; Figure 10). MMI values were somewhat lower in the three lower reaches of Blackwood, which was somewhat surprising given the low level of human development in this area. General had similar low levels of disturbance, but higher MMI values. Unlike General, Blackwood had paved and unpaved roads located near the channel which could have been a source of disturbance. In contrast, Ward also had paved and unpaved roads near the stream channel, but MMI values were among the highest observed. On Trout, one site (near Fountain Place) had an MMI value (22.7) much lower than sites up or downstream. Fountain Place had grazing at one time and is also a low gradient site.

Lowest MMI values were observed in Edgewood (reach 1), Meeks (reach 1), Tallac (reach 1), Taylor (all reaches), and Upper Truckee. Low MMI values in reach 2 of Trout were all associated with small sample sizes less than the 500 count target. One site on General (reach 2) also had very low counts which limits the possible values of the MMI. Some of these sites were in low gradient areas and may indicate the need for different expectations for these habitat types if human disturbance is not present and causing low MMI values. Other sites were very close to the lake shore and may represent more lentic habitats where expectations for taxa richness would naturally be lower.

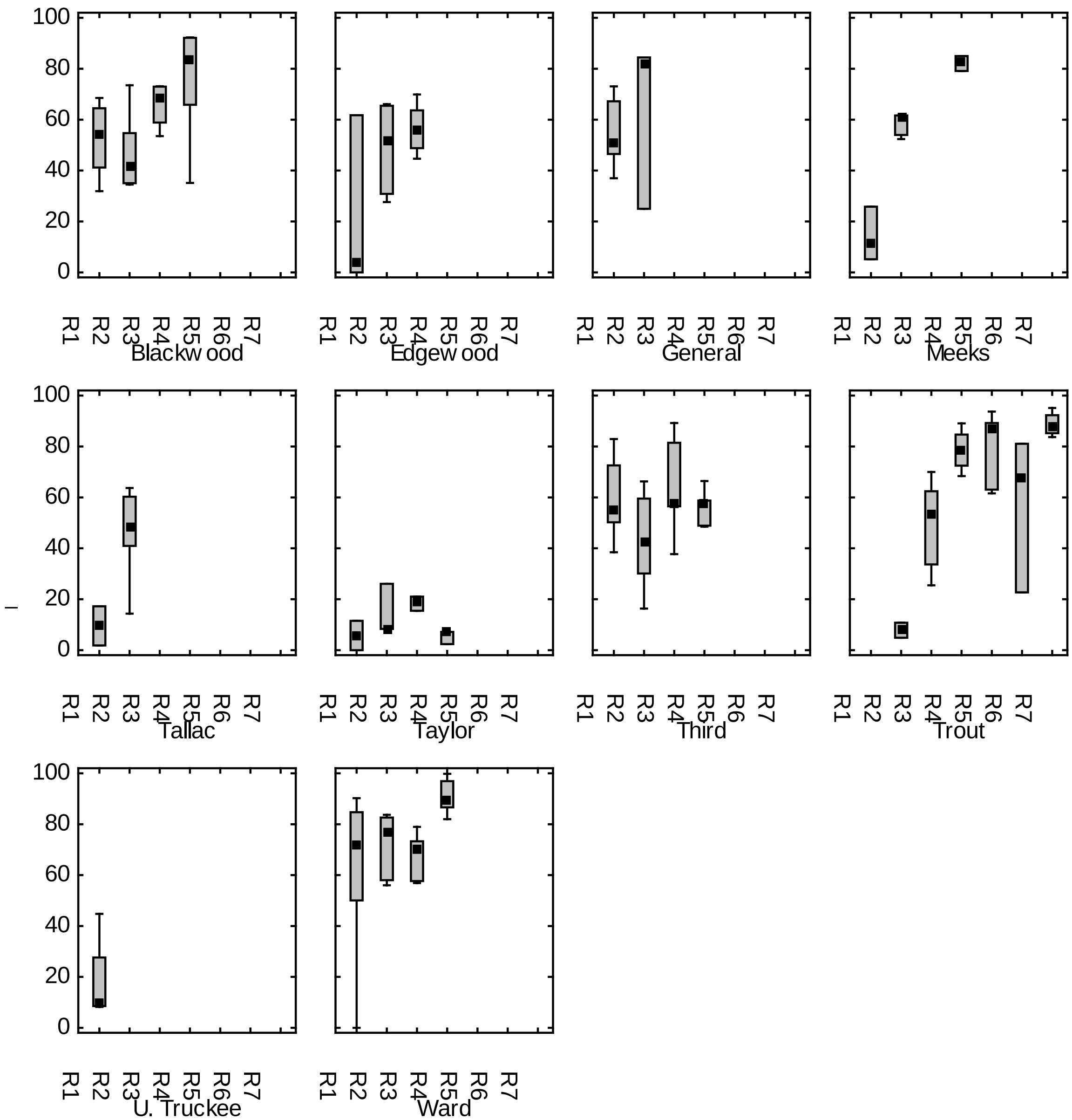

Figure 10. Tahoe Basin MMI values by stream and by reach within a stream. Index values at sites within a reach are summarized by each set of box and whiskers. For each reach are shown the median index values (dark central square), 25th to 75th percentile values (gray boxes), and upper and lower quartiles (whiskers).

DISCUSSION

Metric testing and selection

In general, two rather broad approaches could be taken to the development of a biological assessment protocol for Tahoe Basin streams. One approach would be to take an index developed and tested in a different location and apply it "off the shelf" to Lake Tahoe tributaries. An example of the first approach would be to take the MMI developed for northern California, and simply calculate and apply it to Tahoe Basin streams. At the other end of the spectrum, an alternative approach would be to thoroughly test every possible metric and index and select only those that were the strongest indicators of disturbance for Lake Tahoe tributaries. The western EMAP study is an example of this approach in which an extensive data set was used to test 100's of candidate metrics for correlation with human disturbance (Stoddard et al., 2005).

Rather than choose either of these two approaches, we took a middle path and tested a total of 17 metrics from two multimetric indexes developed and tested for similar geographic areas. These two indexes were selected because the component metrics had been tested against multiple gradients of human disturbance. The analysis was limited to biological measures already tested and vetted by other studies because the data set for Lake Tahoe Basin was somewhat small. With samples from only 10 streams, data were not sufficient to screen or test new or experimental metrics. On the other hand, because the Lake Tahoe Basin represents a unique geographic area, metric testing for disturbance types specific to this area ensured that the resulting index was scientifically relevant and defensible should it be challenged.

Of the 17 metrics tested here, 6 of the 10 metrics used to monitor streams in the Pacific northwest were highly correlated with disturbance. Of the 7 metrics used in the northern California index, only EPT taxa richness and percent non-insect taxa were correlated with site disturbance. On the surface, one might conclude that Tahoe Basin biology is more similar to Oregon and Washington than to northern California. However, a closer look at the metric testing and selection process for the California index reveals that the metrics that correlated with disturbance in Tahoe Basin also correlated with disturbance in northern California, but several of these metrics were not included in the northern California MMI in favor of a more diverse set of metrics. For the California MMI, metrics such as stonefly, caddisfly, and intolerant taxa richness, though correlated with disturbance, were excluded due to their redundancy with EPT richness. Other metrics were selected to capture the range of conditions observed in northern California. Herbst and Silldorff (2006) developed a multimetric index for data collected from the eastern slopes of the Sierra Nevada. They identified 12 core metrics based on their ability to distinguish reference and test sites defined according to road density upstream, bank erosion, and pollution sources upstream. Of these 12, 5 were also correlated with site disturbance in the Tahoe Basin

(total taxa, stonefly, caddisfly, intolerant, and EPT richness). A multimetric index applied to the Truckee River also included several of the same metrics selected for the Tahoe Basin (Tetra Tech, 2004). Thus, diverse methods selected many of the same indicators of biological condition.

For the Tahoe Basin MMI, the EPT taxa were kept separate for their potential ability to diagnose different stressors (Yoder and DeShon, 2002). An increase in mayfly taxa was associated here and in other studies with nutrient enrichment (Miltner and Rankin, 1998). Many stonefly taxa are shredders and their decline is associated with loss of riparian canopy. Caddisfly taxa are very diverse in their habitat needs and feeding ecology and indicate the variety and complexity of habitat available in a stream (Kilbane and Holomuzki, 2004).

For the metrics included in other indexes that were not correlated with site disturbance in the Tahoe Basin, values for these metrics indicated only good to excellent condition for Tahoe Basin streams, sites with values indicating degraded condition for these metrics were not found in the Basin. Similarly, O/E values indicated that many sites in the Tahoe Basin represented reference condition for California. Thus, the range of degradation we are trying to detect in the Tahoe Basin is much more narrow than the range of conditions observed for northern California streams where sources of disturbance were much more diverse and included agriculture, point source pollutants, mining, and heavy timber harvest (Rehn et al. 2005a).

The challenge of quantifying human disturbance

When relating biological condition to human disturbance, most of the error of the correlation is typically due to the inaccuracy associated with quantifying human impact on water resources. The relative amount of stream damage is difficult to attribute to different types of land use (Morley and Karr, 2002). At the outset of this project, our intention was to evaluate the utility of remotely sensed data to estimate human disturbance at the site scale and to use remotely sensed data to compare the influence of human disturbance at multiple spatial scales within the watershed. Although satellite estimates of human disturbance were both extremely accurate and highly precise, the biological metrics were much more strongly correlated with a somewhat crude measure of site condition based on the combination of a single measure of water chemistry (conductivity), riparian condition (large woody debris, bank erosion, and canopy cover), and presence of human structures (road crossings, channel modifications, and buildings).

The measures included in the site disturbance index were certainly not ideal, but represented the best data available to summarize site condition. Percent development, road crossings, and human structures are clearly indicators of human disturbance. In contrast, bank erosion, presence of large woody debris, and canopy cover can be related to natural processes as well as a result of human activities such as logging and clearing of land (Johnson et al., 2003; Martel et al., 2007). Conductivity was not an ideal indicator either. Conductivity measures the

concentration of dissolved ionized solids suspended in the water which may be related to soil erosion, fertilizer, or de-icing chemicals. Although indicative of human disturbance, conductivity values were generally low for Tahoe Basin streams (Potapova and Charles, 2003).

The poor agreement between the estimate of percent development at a very small scale (30 x 50 m) derived from remotely sensed data and measure of site disturbance derived from observations at the site emphasizes the difficulty in using reflectance measures to summarize the many and complex ways that human land use degrades water resources. Thus, at a larger spatial scale, when comparing one stream to another, the amount of developed area provides a reasonable approximation of watershed condition (Booth et al., 2004). But for comparisons at a smaller spatial scale, that is, sites within Tahoe Basin watersheds, a more accurate measure was needed that summarized the different aspects of human disturbance. As an illustration of the interplay between watershed and site scale disturbance, Karr (2006) found that the highest percentage of development was the best predictor of biological condition, irrespective of spatial scale.

Metrics were also correlated with elevation

Human disturbance is naturally associated with elevation and slope around Lake Tahoe, people build closer to the Lake on flatter ground. The concern for this study was that elevation and human influence would be too confounded to isolate metric response to human disturbance. A fundamental requirement for a bioassessment protocol is that it be sensitive to human disturbance but immune to natural variability. Two lines of evidence suggest that biological metrics were primarily associated with human disturbance rather than elevation. First, the MMI was only correlated with elevation at lower elevations, where most of the human disturbance occurs. At higher elevations, MMI was not correlated with elevation. Second, site condition and elevation were not highly correlated: a range of site conditions were observed at most elevations. An exception was high elevation sites (>7200 ft) for which minimal site disturbance was observed. Multiple regression analysis showed significant correlations with *both* elevation and site condition for all biological metrics included in the Tahoe Basin MMI. The correlation of the MMI and its component metrics with elevation more likely reflects a summary of cumulative effects of disturbance at the watershed scale not captured by the measure of site disturbance. Additional sampling at low elevation sites with minimal human influence could resolve this question. If high MMI values are observed at sites closer to the lake with minimal disturbance and lower gradient, then concerns about elevation as a confounding natural variable can be eliminated.

Detecting change through time

Replicate samples were not collected for this study; therefore, estimates of variance could not be derived. Variance associated with replicate sampling within a year and across years is needed to define the optimal trend monitoring design (Kincaid et al., 2004). Varying the number of samples collected each year or the number of replicate samples collected at each location influences the amount of change in biological condition (as measured by MMI) that can be detected through time. Samples collected from neighboring sites or sites within reaches could have been used to calculate variance. Given the high correlation between MMI and site condition, this approach was not pursued because neighboring samples did not necessarily represent replicate conditions.

Although variance estimates were not available for the 2003 MMI values, results for the Pacific northwest B-IBI are likely to be similar. Based on repeat visits to 40 stream sites, the B-IBI could detect a 10% (2% per year) change in biological condition over a five year period with a statistical confidence of 80% (Fore, unpublished data).

RECOMMENDATIONS

The development of a biological monitoring program can be a complex and iterative process as the development of one element can trigger modifications to another. Working from a general framework that identifies the specific steps of a successful monitoring program supports this process by identifying the elements that are complete and those which remain to be done. For the Tahoe Basin, the 5 steps identified by Ward et al. (1986) help guide the process toward an overall monitoring strategy for the biological condition of streams (Table 8). This systematic approach to developing a monitoring program supports the foundations of biological monitoring identified by Karr (2006a) and interpreted for the Tahoe Basin (Karr, 2006b).

Table 8. Steps in the design of a biological monitoring program (modified from Ward et al., 1986), specific items within each step, an example or comment linking the step or item to the Tahoe Basin, and whether an item has been completed or is addressed by the current document.

Step	Tahoe example or comment	Complete?	This document
Step 1: Identify information expectations	Utilize a consensus approach among citizens and stakeholders to articulate desirable future conditions for Lake Tahoe Basin		
Resource condition goals	Identify indicators of environmental conditions	Yes	
Management goals	Agree upon measurable standards for the indicators, e.g., relative to reference condition	Partial	
Monitoring goals (as statistical hypotheses)	Set targets for management actions (e.g., “At least 50% of stream miles within the Basin should represent reference condition. If target condition not met, stream restoration projects will be implemented”)		Partial
Step 2: Develop statistical design criteria	Framework to translate goals into a survey design		
Define ‘population’ to be sampled	Streams sites within the Tahoe Basin	Yes	
Evaluate variability of protocol	Determine how many stream samples are needed to be 90% confident that streams are meeting their target condition	No	
Select appropriate statistical survey method	Probabilistic sampling of all sites within the Basin (sensu EPA, 2006b)		Yes

Step	Tahoe example or comment	Complete?	This document
Select appropriate statistical test	Binomial model to estimate confidence intervals for % stream miles that match reference condition; linear regression of MMI vs. year for trend assessment		Yes
Step 3: Develop a sampling plan	Specifics of how to collect data and a list of which sites are visited each year		
Identify what to measure	Macroinvertebrate assemblages	Yes	
Identify where & when to sample	Riffle habitats during the late summer/early fall	Yes	
Define how frequently to sample	Repeat visits to 10 sites every 5 y for trend monitoring; 30 new sites every year for status assessment		Yes
Step 4: Develop written protocols	Detailed description of procedures that other could follow		
Data collection, sample preservation, and transport	Details of field equipment needed to collect macroinvertebrates and convey them to the lab for identification	Yes	
Lab analysis	Identify a minimum of 500 individuals; calculate Tahoe Basin MMI using rules outlined in this document		Yes
Data management	Archive data in a database with metric calculation rules automated		Yes
Step 5: Report results	Specify report format, distribution, and peer-review process to determine whether goals of Step 1 were met		
Define report structure, format, and content	See BRMEP	Yes	
Determine frequency and distribution of reports	Every 5 years	Yes	

Further refine information expectations (Step 1)

Some of the elements under Step 1 are outside the scope of this document because they rely on the input and decisions of a larger group of regional managers and citizens (Davies and Jackson, 2006). Complete under this step is the identification of resource condition goals and the indicators that will be used to track environmental condition. The next item needed is an agreement on management targets and specific monitoring goals. In short, a standard must be defined along with numeric criteria that indicate an unacceptable departure from the standard.

EPA guidance recommends a reference condition approach to establishing monitoring goals and standards (EPA 2006a, c; Stoddard et al., 2006). Sites with minimal human influence are selected according to objective criteria, the biological condition at the sites is assessed (e.g., using a multimetric index), and site information is evaluated to determine whether it matches reference condition or departs significantly. The amount of departure from expectation tolerated is not a statistical or scientific decision but is directly related to the level of protection desired by the larger community. For example, a higher standard might be expected for streams in an area such as the Tahoe Basin, along with a higher level of protection.

As an example, the Maryland Biological Stream Survey' monitoring program uses the following objective criteria to define reference condition (EPA, 2006c):

- pH >6
- ANC >50 μeq/L
- dissolved oxygen >4 ppm
- nitrate <300 μeq/L
- urban land use <20% of the catchment area
- forest land use >25% of the catchment area
- remoteness rating = optimal or suboptimal (>10 on 0-20 scale)
- aesthetics rating = optimal or suboptimal
- instream habitat rating = optimal or suboptimal
- riparian buffer width >15 m
- no channelization
- no point source discharges

Once objective criteria are defined for reference condition, the biological indicators measured at those sites represent a benchmark to set expectations for biological condition at sampled sites. Thus, using a probabilistic sampling design (see below), the percentage of stream miles in reference condition (or 1 or 2 standard deviations away from reference condition) could be calculated and reported.

- *Identify reference condition and sample reference sites in different habitat types.*

For Tahoe Basin streams, a similar process is needed to identify reference condition. Criteria for Tahoe Basin might be based on development estimated from remotely sensed data and measures of site condition. Approximately 30 sites should be adequate and many may have already been sampled during 2003.

Additional reference sites should be identified for low elevation, low gradient reaches and sampled to determine if expectations for this habitat type need to be modified. Although the Tahoe Basin MMI was correlated with elevation, high MMI values in the lowest reaches on Ward and General where human disturbance was minimal suggest that low gradient sites can potentially support similar macroinvertebrate communities as those found in upstream sites. If MMI values are similar for reference sites at different elevations and habitat types, then separate criteria would not be needed. If reference sites differ, a simple numeric adjustment to the metric scoring process can be used to adjust the MMI to ensure that index values indicate similar biological condition across the Tahoe Basin. Approximately 10 sites should be adequate to answer this question.

- *Translate monitoring goals into statistical hypotheses.*

Once reference sites are identified, the next task must be to define what values of the Tahoe Basin MMI represent a departure from reference condition (many agencies use the 10^{th} percentile of index values for reference sites with a buffer area below that to account for natural variability). This aspect of the task is largely a quantitative and statistical exercise but cannot replace the decision regarding how much departure from reference condition is acceptable. This decision is derived from societal values rather than numbers. As an example, returning salmon are a cultural icon and economic resource in the Pacific northwest. B-IBI values derived from salmon-bearing streams could be used to set an expected standard for the biological health of regional streams (Karr and Rossano, 2001). One example of a monitoring goal for Tahoe Basin might be “50% of stream miles representing reference condition.”

Continue defining statistical design criteria (Step 2)

- *Define the sample population for Tahoe Basin streams.*

For Tahoe Basin, the population to be sampled can be defined as all tributary streams in the Basin. Exceptions to this definition of the population might include stream sites close to the lake where a stream transitions to lentic habitat. Streams must also have adequate flow and depth to enable sampling. Other sites could be excluded from the population if they are of less concern, for example, sites within protected areas or sites that are outside the management purview of the agency.

Within the population, sampling units must also be defined. Sampling units are the elements of the population that are actually measured. For medical surveys, a person is typically the sampling unit. In contrast, rivers and streams are continuous resources and do not form simple, discrete sampling units; therefore, some method must be used to delineate stream sections into sampling units (Stevens and Olsen, 2003). Many of these issues have been resolved and specific software has been developed to randomly select sampling locations from continuous resources such as rivers, estuaries, and reefs (EPA, 2006b).

- *Revisit stream sites to estimate variance of MMI.*

The variability of the stream sampling protocol for macroinvertebrates should be evaluated using repeat visits to a site. Repeat visits are needed to estimate the natural variability associated with the Tahoe Basin MMI. Repeat visits should be made to 30–40 stream sites that were previously sampled in 2003. Stream sites should be sampled twice during the typical sampling period, sampling events should be on different days, and macroinvertebrate samples should be collected by different crews to capture the various sources of measurement error. Variance estimates for the MMI derived from the repeat visits can be used in statistical power analysis to compare different types of trend models (e.g., that vary by number of visits or years) to identify the most efficient trend monitoring design. Variance can also be used to predict the amount of change that the MMI could potentially detect for different numbers of replicate samples collected before and after a stream restoration project, for example.

- *Use a regression model for trend monitoring and paired testing for targeted monitoring.*

Selecting the appropriate statistical test depends on the question being asked. In general, two situations are most common in the context of resource monitoring. For trend monitoring, sites are randomly selected from a pool of all possible sites within the region of interest. Sampling these sites in one year yields estimates of the status of the resource; sampling these sites through time yields an estimate of the trend in condition to assess whether stream sites are improving or declining through time. The appropriate statistical model for this situation is a regression analysis of the MMI against time (Urquhart et al., 1998; Larsen et al., 2001).

The second situation involves a specific known site, or set of sites, that need to be evaluated or compared. Examples of this situation include 1) comparison of sites upstream and downstream of a restoration project or 2) sites evaluated for their condition before and after best management practices have been applied. This situation is an example of ‘targeted monitoring.’ For these scenarios, a two-sample test would be an appropriate statistical test (Peterman, 1990).

Develop a sampling plan (Step 3)

A sampling plan connects the general monitoring questions to the exact specific details of *what*, *where*, and *when.* A formal sampling plan insures that the correct information is obtained to

answer the questions being asked (Urquhart et al., 1998; Olsen et al., 1999; Yoccoz et al., 2001). For Tahoe Basin streams, the *what* will be benthic macroinvertebrates from riffle habitat. The *where* and *when* of sampling depends on the types of questions that are being asked.

In general, monitoring questions can be divided into three types related to 1) assessing the current *status* of the resource, 2) detecting *trends* over time, and 3) evaluating conditions at specific, *targeted* locations. Status, trend, and targeted monitoring all differ in the manner in which sampling units (e.g., stream sites) are selected from among the total population of all possible sampling units (Stevens and Olsen, 1999). Status monitoring is best accomplished with random selection of sampling locations every year. The best sampling designs for detecting trend initially select sampling locations randomly, but then re-visit the same locations either every year or in a rotating pattern every fourth or fifth year (Larsen et al., 2001). For targeted monitoring, locations are selected based on specific criteria designed to answer a particular question, such as sites with best management practices in place or sites with known sources of disturbance. The best approach for balancing multiple monitoring goals is to first identify the total number of samples that can be collected each year and then assign a percentage of the sampling effort to each monitoring goal according to its relative importance.

Status monitoring

For assessing the overall status of a resource, such as streams in the Tahoe Basin, the best approach randomly selects as many sampling units from the population as possible. The primary advantage of random site selection is that any summary statistics derived from a random sample will be unbiased for the entire population, including all the sites that were *not sampled*. Although site-specific information about unsampled sites is not available, the general overall condition of sites not sampled can be inferred from the sampled sites. Evaluating sites without the expense of sampling is certainly cost-effective. This is the rationale behind the probabilistic survey sampling designs of EPA's Environmental Monitoring and Assessment Program (EMAP; Larsen, 1997; Paulsen et al., 1998). In addition to an unbiased estimate of overall condition, the uncertainty, or variability, associated with any estimate of stream condition can be derived for the entire population (Paulsen et al., 1998; Urquhart et al., 1998; Yoccoz et al., 2001).

Status monitoring is designed to answer questions such as:

What percentage of stream miles in the Tahoe Basin represent "reference" condition?

What percentage of stream miles within one mile of Lake Tahoe are characterized by degraded biological condition?

Trend monitoring

The goal of trend monitoring is to detect a change through time should a change occur. The recommended statistical model for trend detection is a regression model in which the variable of interest, e.g., an index of stream condition, is regressed against time. Sampling designs in which the same sites are visited each year have the greatest power to detect a trend because each site is effectively compared with itself through time (Larsen et al., 1995; Urquhart et al., 1998). In contrast, selecting a new random set of sites each year introduces a large amount of variability due to site differences and makes it more difficult to detect trend. The same sites may be visited each year, or visited in a rotating panel design, such that each site is visited every fourth or fifth year (Skalski, 1990).

Trend monitoring is designed to answer questions such as:

Is the biological condition of Tahoe Basin streams improving or declining?

Is decline (or improvement) in biological condition associated with different types of land use?

Targeted monitoring

The utility of a targeted monitoring approach can be imagined for a number of scenarios: to determine whether management practices are effective, to regulate specific point sources, or to refine data collection methods. The caveat associated with selecting sites non-randomly for whatever reason is that any conclusions related to environmental conditions at those sites will *only apply to those sites*, not to sites in general (Olsen et al., 1999; Hughes et al., 2000). No matter how statistically significant the trend or change observed for those sites, the conclusions cannot be extended to other sites with any known level of confidence.

Targeted monitoring may be used to answer questions such as:

Does removal of box culverts improve the biological condition of a stream reach?

What types of stream restoration are most effective for improving stream condition?

Under Step 3, what remains for Tahoe Basin is the adoption of a probabilistic sampling plan that randomly selects stream sites for sampling. The easiest solution here is to use the software developed by EPA for this purpose (EPA, 2006b). To implement this type of sampling plan, decisions remain regarding the amount of sampling to be accomplished each year along with the relative allocation of sampling effort to the three types of monitoring.

- *Determine how many sites can be visited each year.*

The simplest approach to developing a sampling plan is to start with the number of sites that can be visited each year and work backwards to allocate sampling sites to program goals according to their relative priority.

- *Allocate sampling effort to status, trend, and targeted monitoring according to program needs.*

The allocation of sampling effort devoted to status, trend, and targeted monitoring should be objectively determined according to program needs and priorities. To illustrate this process, suppose that Tahoe Basin agencies decide to allocate sampling effort approximately equally to status, trend, and targeted monitoring and that 50 macroinvertebrate stream samples can be collected each year. For status monitoring, 15 sites would be sampled each year (Table 9). For large regional areas, sites are typically grouped into 4 or 5 geographic areas to minimize driving time and a different area sampled within a 5-year rotation pattern. Given the relatively small size of Tahoe Basin, grouping sites by proximity may be less important. If so, then a new set of status sites would be selected each year from the entire Basin.

For trend monitoring, 15 sites would be randomly selected during the first year of sampling and then revisited either the following year or during a subsequent year. If a 5-year rotation is selected, the same group of sites must be sampled together every fifth year. Sites should not stray from one group to another.

For targeted monitoring, sample sites will likely change from year to year as projects develop and come to an end. Targeted monitoring of restoration projects may require a greater statistical precision to detect small changes in a shorter period of time; therefore, replicate sampling is assumed for these sites. The total sampling effort each year would include visits to approximately 36 locations and laboratory processing of approximately 48 samples. The example described here represents only one possibility, the best sampling plan for the Tahoe Basin may be quite different.

Some flexibility is possible for the status and targeted monitoring programs in that more or fewer sites may be sampled in a given year. In contrast, the trend sites are not flexible. Once trend sites are selected within a region, they must all be sampled during their rotation year. The statistical model (MMI for sites regressed against year) assumes that the same sites are sampled during each time period which could be every year, every other year, or every fifth year according to program needs. If a site is not sampled in a particular year when it was originally planned to be sampled, data collected from that site for *all years previously sampled* must be removed from the analysis. Sites cannot be substituted or replaced because the trend model works by comparing each site to itself through time. Comparing the same sites through time is the most sensitive model for detecting change.

Table 9. Example sampling plan to monitor status, trend, and targeted sites. Cell values represent the number of sites x the number of replicate samples. Five geographic areas are defined (A-E) and one is visited each year. Every year, 15 new sites are randomly selected for status monitoring. For trend monitoring, 15 sites are randomly selected in only the first year of sampling; in subsequent years the same sites are visited in rotation. Targeted sites are reserved for special projects and assigned according to specific management needs each year.

Sampling purpose	Panel	Year 1	Year 2	Year 3	Year 4	Year 5	*Etc.*
Status	**A**	15 x 1					
	B		15 x 1				
	C			15 x 1			…
	D				15 x 1		
	E					15 x 1	
Trend	**A**	15 x 1					
	B		15 x 1				
	C			15 x 1			…
	D				15 x 1		
	E					15 x 1	
Target		6 x 3	6 x 3	6 x 3	6 x 3	6 x 3	…
Total		36 sites, 48 samples	36 sites, 48 samples	36 sites, 48 samples	36 sites, 48 samples	36 sites, 48 samples	

Develop written protocols (Step 4)

Under Step 4, the goal is to provide sufficient details for all procedures such that future personnel can collect comparable data. For status and trend monitoring, the time frame can be long (>20 years) and in order to make meaningful comparisons through time the same methods must be used to collect and analyze the data. If changes are made, they must be made consciously and must be accompanied by some type of analysis equating the results from the old method to the new method so that data from the past are not orphaned.

For Tahoe Basin streams, field operations went smoothly during 2003 and crews were able to successfully obtain macroinvertebrate data. This document outlines the method for summarizing macroinvertebrate data using the Tahoe Basin MMI. What remains under this step relates to data management.

- *Use data management software to calculate index values and store data.*

Repetitive calculations performed with long time periods in between invite error. The data used in this project were delivered from as individual files from Access and Excel software. An integrated approach to data management is needed with error checking each year and a single location (file) for all the data collected during multiple years. Calculation of the MMI should be programmed and applied uniformly to every data set. Statistical calculations related to the percentage of stream miles in different condition classes could also be coded and calculated automatically each year. Lack of attention to data management is one of the greatest underlying threats to the success of a long-term monitoring program (Fore, 2003).

Report results (Step 5)

Over the years, monitoring results have typically been reported in written documents. As a consequence, the time between data collection and distribution of the results can be long. If the monitoring program is long-term, a new report is not typically needed each year and publication of the data may have to wait for the next report. Most monitoring programs generate new data every year, and increasingly more people are interested in this information. To meet this need and ease the burden associated with writing reports, many agencies are moving toward web-based delivery of data products.

- *Report monitoring results dynamically using web technology.*

Given the large number of agencies interested in the condition of Tahoe Basin streams, web technology could simplify reporting procedures. The type of information typically found in a monitoring report could be written once and posted as web pages. The tables and figures in the report could be updated as soon as new data become available without rewriting or redistributing the report. The USGS stream gauge system is an example of a real-time distribution of data (USGS, 2007). Oregon and Florida are also implementing similar data systems for biological assessment data (FDEP, 2004; ODFW, 2007). A more integrative approach developed by EPA allows a web client to browse the data and calculate statistics in real-time, e.g., the percentage of stream miles with index values below a specified value (EPA, 2006d).

CONCLUSIONS

Changes to the resident macroinvertebrate community were highly correlated with an independent measure of human disturbance for Tahoe Basin streams. Results from this study parallel the patterns observed in other regional assessments of biological response to disturbance in northern California, Washington, Oregon, and other western states. The Tahoe Basin MMI summarizes the biological changes observed as disturbance increased and can be used at the basin level to assess status of stream condition, track trends through time, and evaluate specific changes in biological condition associated with stream restoration or implementation of best management practices.

Karr (2006a) describes seven foundations of biological monitoring that are relevant for the protection of aquatic living systems. In a separate report these foundations were applied to biological monitoring in the Tahoe Basin (Karr, 2006b). Many of the elements that he identifies are being actively addressed in the Tahoe Basin for stream monitoring and other elements are addressed specifically by this report. For example, management agencies in Tahoe Basin recognize that Tahoe Basin streams have been altered by human disturbance (Karr's Foundation 1) and that legislative and regulatory mandates can be applied to reverse this trend (Foundation 2). Furthermore, Tahoe Basin agencies recognize the inherent value of biological indicators for integrative monitoring of ecological condition (Foundations 3 and 4). This report addressed the specific elements needed to translate biological data into a bioassessment tool (Foundation 5) and implement a long-term monitoring program to assess stream condition (Foundation 6). The ultimate success of the monitoring program relies on completing the cycle which means communicating the results to the people asking the questions about resource condition (Foundation 7).

In order to implement a formal sampling plan for the Tahoe Basin, a few items remain to be addressed such as the development of a definition of reference condition for Tahoe Basin, assessment of reference sites in low elevation and low gradient habitats to calibrate the index, and replicate sampling to estimate index variance. Tahoe Basin management agencies are moving away from pilot studies toward long-term monitoring programs; as a consequence, any future data collection efforts should be carefully matched to specific information needs. The MMI was specifically developed and calibrated for Lake Tahoe tributaries in order to support the management goals of local agencies for the protection of Lake Tahoe Basin.

REFERENCES

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates, and fish. 2nd ed. EPA 841-B-99-002. Washington, D. C. Office of Water. www.epa.gov/owow/monitoring/rbp/index.html.

Booth, D. B., J. R. Karr, S. Schauman, C, P. Konrad, S. A. Morley, M. G. Larson, and S. J. Burges. 2004. Reviving urban streams: land use, hydrology, biology, and human behavior. Journal of the American Water Resources Association 40(5):1351-1364.

CDFG (California Department of Fish and Game). 2006. CalEDAS Bioassessment Database www.dfg.ca.gov/cabw/cabwhome.html.

Carter, J. L. and V. H. Resh. 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society 20:658-682.

Davies, S. P. and S. K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic systems. Ecological Applications 16:1251-1266.

EPA (U.S. Environmental Protection Agency). 2002. Summary of Biological Assessment Programs and Biocriteria Development for States, Tribes, Territories, and Interstate Commissions: Streams and Wadeable Rivers. EPA-822-R-02-048. Washington, D. C. Office of Environmental Information and Office of Water. www.epa.gov/bioindicators/html/program_summary.html.

———. 2006a Draft. Consolidated Assessment and Listing Methodology – Toward a compendium of best practices. Washington, D. C. Office of Wetlands, Oceans, and Watersheds. www.epa.gov/owow/monitoring/calm.html.

———. 2006b. Aquatic Resources Monitoring. www.epa.gov/nheerl/arm/.htm.

———. 2006c. Best Practices for Identifying Reference Condition in Mid-Atlantic Streams. Washington, D. C.: Office of Environmental Information. EPA-260-F-06-002. www.epa.gov/bioindicators/html/publications.html

———. 2006d. Data browser for Western EMAP. http://www.epa.gov/bioindicators/statprimer/data_browser.html

FDEP (Florida Department of Environmental Protection). 2004. Bioassessment Ecosummaries. http://water.dep.state.fl.us/eswizard/eco_query.asp

Feck, J. and R. O. Hall. 2004. Response of American dippers (*Cinclus mexicanus*) to variation in stream water quality. Freshwater biology 49:1123-1137.

Fore, L. S. 2003. Developing biological indicators: Lessons learned from mid-Atlantic streams. EPA 903/R-003/003. Ft. Meade, Maryland: Office of Environmental Information and Mid-Atlantic Integrated Assessment Program, Region 3. Available: http://www.epa.gov/bioindicators.

Fore, L. S., J. R. Karr, and R. W. Wisseman. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15:212-231.

Fore, L. S., K. Paulsen, and K. O'Laughlin. 2001. Assessing the Performance of Volunteers in Monitoring Streams. Freshwater Biology 46: 109-123.

Frazier, J. W., K. B. Roby, J. A. Boberg, K. Kenfield, J. B. Reiner, D. L. Azuma, J. L. Furnish, B. P. Staab, S. L. Grant. 2005. Stream Condition Inventory Technical Guide. USDA Forest Service, Pacific Southwest Region - Ecosystem Conservation Staff. Vallejo, CA.

Johnson, L. B., D. H. Breneman, and C. Richards. 2003. Macroinvertebrate community structure and function associated with large wood in low gradient streams. River Research and Applications 19: 199-218.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications 10:1456–1477.

Herbst, D. B. and E. L. Silldorff. 2006. Comparison of the performance of different bioassessment methods: similar evaluations of biotic integrity from separate programs and procedures. Journal of the North American Benthological Society 25: 513–530

Hodkinson, I. D. and J. K. Jackson. 2005. Terrestrial and aquatic invertebrates as bioindicators for environmental monitoring, with particular reference to mountain ecosystems. Environmental Management 35:649–666.

Hughes, R. M., Paulsen, S. G., Stoddard, J. L. 2000. EMAP-Surface Waters: a multiassemblage, probability survey of ecological integrity in the USA. Hydrobiologia 422:429-443.

Karr J. R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications 1:66-84.

Karr, J. R. 1998. Rivers as sentinels: using the biology of rivers to guide landscape management. River Ecology and Management: Lessons from the Pacific Coastal Ecosystem (Eds. R. J. Naiman & R. E. Bilby), pp. 502-528. Springer, NY.

Karr, J. R. 2006a. Seven foundations of biological monitoring and assessment. Biologia Ambientale 20(2): 7-18.

Karr, J. R. 2006b. Indicator Development in the Tahoe Basin: Review of Progress and Thoughts on Needs for the Future. Report prepared for USFS Tahoe Basin Management Unit.

Karr, J. R., and E. W. Chu. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Washington, DC: Island Press.

Karr, J. R. and E. M. Rossano. 2001. Applying public health lessons to protect river health. Ecology and Civil Engineering 4: 3−18.

Kerans B. L. and Karr J. R. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4:768-785.

Kilbane, G. M. and J. R. Holomuzki. 2004. Spatial attributes, scale, and species traits determine caddisfly distributional responses to flooding. Journal of the North American Benthological Society 23:480-493.

Kincaid, T. M., D. P. Larsen, and S. N. Urquhart. 2004. The structure of variation and its influence on the estimation of status: Indicators of condition of lakes in the northeast, U.S.A. Environmental Monitoring and Assessment 98:1-21.

Larsen, D. P. 1997. Sample survey design issues for bioassessment of inland aquatic ecosystems. Human Ecol. Risk Assess. 3: 979-991.

Larsen, D. P. and A. T. Herlihy. 1998. The dilemma of sampling streams for macroinvertebrate richness. Journal of the North American Benthological Society 17:359-366.

Larsen, D. P., T. M. Kincaid, S. E. Jacobs, and N. S. Urquhart. 2001. Designs for evaluating local and regional scale trends. BioScience 12:1069-1078.

Larsen, D. P., N. S. Urquhart, and D. L. Kugler. 1995. Regional scale trend monitoring of indicators of trophic condition of lakes. Water Resources Bulletin 31: 117-140.

Martel, N. M. A. Rodriguez, and P. Berube. 2007. Multi-scale analysis of responses of stream macrobenthos to forestry activities and environmental context. Freshwater Biology 52:85-97.

Merritt, R. W., and K. W. Cummins, eds. 1996. An introduction to the aquatic insects of North America, 3rd ed. Dubuque, Iowa: Kendall/Hunt Publishing Company.

Miltner, R. J., and E. T. Rankin. 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshwater Biology 40:145–158.

Morley, S. A., and J. R. Karr. 2002. Assessing and restoring the health of urban streams in the Puget Sound Basin. Conservation Biology 16: 1498–1509.

Nakano, S. and M. Murakami. 2001. Reciprocal subsidies: dynamic interdependence between terrestrial and aquatic food webs. Proceedings of the National Academy of Sciences 98:166-170.

Ode, P. R. and A. C. Rehn. 2005. Probabilistic assessment of the biotic condition of perennial streams and rivers in California. Report to the State Water Resources Control Board. California Department of Fish and Game, Aquatic Bioassessment Laboratory, Rancho Cordova, CA.

ODFW (Oregon Department of Fish and Wildlife). 2007. Coastal Coho Assessment. http://nrimp.dfw.state.or.us/Oregonplan/default.aspx?p=1.

Olsen, A. R., Sedransk, J., Edwards, D., Gotway, C. A., Liggett, W., Rathbun, S., Reckhow, K.H. and Young, L. J. 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54: 1-45

Paulsen, S. G., R. M. Hughes, and D. P. Larsen. 1998. Critical elements in describing and understanding our nation's aquatic resources. Journal of the American Water Resources Association 34:995-1005.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47: 2-15.

Potapova, M. and D. F. Charles 2003. Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition Freshwater Biology 48:1311–1328.

Rehn, A. C., P. R. Ode, C. P. Hawkins. 2007. Comparisons of targeted riffle and reach-wide benthic macroinvertebrate samples: implications for data sharing in stream condition assessments. Journal of the North American Benthological Society, in press.

Rehn, A. C., P. R. Ode, and J.T. May. 2005a. Development of a benthic index of biotic integrity (B-IBI) for wadeable streams in northern coastal California and its application to regional 305(b) reporting. Unpublished technical report for the California State Water Quality Control Board. www.swrcb.ca.gov/swamp/docs/northc1.pdf.

Rehn, A. C., P. R. Ode, J. T. May. 2005b. A quantitative tool for assessing the integrity of southern coastal California streams. Environmental Management 35: 493–504.

Sabo, J. L. and M. E. Power. 2002. River-watershed exchange: effects of riverine subsidies on riparian lizards and their terrestrial prey. Ecology 83:1860-1869.

Simon, T. P. (Ed.). 2002. Biological Response Signatures: Indicator Patterns Using Aquatic Communities. CRC Press, Boca Raton, FL.

Skalski, J. R. 1990. A design for long-term status and trends monitoring. Journal of Environmental Management 30:139-144.

Stevens, D. L., Jr. and Olsen, A.R. 1999. Spatially restricted surveys over time for aquatic resources. Journal of Agricultural, Biological, and Environmental Statistics, 4, 415-28.

Stevens, D. L., Jr. and Olsen, A. R. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14:593-610.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectation for the ecological condition of streams: the concept of reference condition. Ecological Applications, 16(4), 2006, pp. 1267–1276.

Stoddard, J. L., D. V. Peck, A. R. Olsen, D. P. Larsen, J. Van Sickle, C. P. Hawkins, R. M. Hughes, T. R. Whittier, G. Lomnicky, A. T. Herlihy, P. R. Kaufmann, S. A. Peterson, P. L. Ringold, S. G. Paulsen, R. Blair. 2005. Environmental Monitoring and Assessment Program (EMAP) Western Streams and Rivers: Statistical Summary. US EPA, Office of Research and Development, Washington, DC. EPA 620/R-05/006.

Tetra Tech, Inc. 2004. Biological Condition Index Development for the Truckee River: Benthic Macroinvertebrate Assemblage. Final report to Nevada Division of Environmental Protection.

TIIMS (Tahoe Integrated Information Management System). 2007. Pathway 2007. www.pathway2007.org/.

USGS (U.S. Geological Survey). 2007. Real-time water data for USA. waterdata.usgs.gov/nwis/rt.

Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. Ecological Applications. 8:246-257.

Vowell, J. L. 2001. Using stream bioassessment to monitor best management practice effectiveness. Forest Ecology and Management 143:237-244.

Walsh, C. J. 1997. A multivariate method for determining optimal subsample size in the analysis of macroinvertebrate samples. Marine and Freshwater Research 48(3), 241-8.

Ward, R. C., Loftis, J. C., and G. B. McBride. 1986. The “data rich but information poor” syndrome in water quality monitoring. Environmental Management 10:291-297.

Winter, J. G., K. M. Somers, P. J. Dillon, C. Paterson, and R. A. Reid. 2002. Impacts of golf courses on macroinvertebrate community structure in Precambrian shield streams. Journal of Environmental Quality 31:2015-2025.

Yoccoz, N. G., Nichols, J. D. and Boulinier, T. 2001. Monitoring of biological diversity in space and time. Trends In Ecology & Evolution 16: 446-453.

Yoder, C.O. and J.E. DeShon. 2002. Using biological response signatures within a framework of multiple indicators to assess and diagnose causes and sources of impairments to aquatic assemblages in selected Ohio Rivers and streams. Pp. 23-82, in T. P. Simon (ed.) Biological Response Signatures: Indicator Patterns Using Aquatic Communities. CRC Press, Boca Raton, FL.

APPENDIX A. DESCRIPTION OF METRICS INCLUDED IN THE B-IBI

Ten metrics are included in the benthic index of biotic integrity (B-IBI) that was developed and tested in streams in the Pacific northwest and Japan (Karr, 1998).

Total taxa richness. The biodiversity of a stream declines as flow regimes are altered, habitat is lost, chemicals are introduced, energy cycles are disrupted, and alien taxa become dominant. Total taxa richness includes all the different invertebrates collected from a stream site: mayflies, caddisflies, stoneflies, true flies, midges, clams, snails, and worms.

Mayfly (Ephemeroptera) taxa richness. The diversity of mayflies declines in response to most types of human influence. Many mayflies graze on algae and are particularly sensitive to chemical pollutants (e.g., from mine tailings) that interferes with their food source. Mayflies may disappear when heavy metal concentrations are high while caddisflies and stoneflies are unaffected. In nutrient-poor streams, manure and fertilizers can increase the numbers and types of mayflies present. If many different taxa of mayflies are found while the variety of stoneflies and caddisflies is low, enrichment may be the cause.

Stonefly (Plecoptera) taxa richness. Stoneflies are the first to disappear from a stream as human disturbance increases. Many stoneflies are predators that stalk their prey and hide around and between rocks. Hiding places between rocks are lost as sediment washes into a stream. Many stoneflies are shredders and feed on leaf litter that drops from an overhanging tree canopy. Most stoneflies, require cool water temperatures and high oxygen to complete their life cycles.

Caddisfly (Trichoptera) taxa richness. Different caddisfly species feed in a variety of ways: Net-spinners trap food, grazers scrape algae from the tops of exposed rocks, and predators hunt live prey. Many caddisflies build gravel or wood cases to protect them from predators; others are predators themselves. Even though they are very diverse in habit, taxa richness of caddisflies declines steadily as humans eliminate the variety and complexity of their stream habitat.

Intolerant taxa richness. Animals identified as intolerant are the most sensitive taxa; they represent approximately 5-10% of the taxa present in the region. These animals are the first to disappear as human disturbance increases.

Clinger taxa richness. Taxa defined as clingers have physical adaptations such as ventral suckers, dorsoventral flattening, well-developed tarsal claws, or construct retreats that they attach to the substrate; thus, they are able to “cling” to smooth substrates in fast water. These animals require open areas between rocks and cobble along the bottom of the stream; consequently, they are particularly sensitive to fine sediments that fill these spaces and eliminate the variety and complexity of small habitats. Clingers may use these areas to forage, escape from predators, or

lay their eggs. Sediment also prevents clingers from moving down deeper into the stream bed, or hyporheos, of the channel.

Long-lived (semi-voltine) taxa richness. These invertebrates require more than one year to complete their life cycles; thus, they are exposed to all the human activities that influence the stream throughout one or more years. If the stream is dry part of the year or subject to flooding, these animals may disappear. Loss of long-lived taxa may also indicate an on-going problem that repeatedly interrupts their life cycles.

Percent tolerant. Tolerant animals are present at most stream sites, but as disturbance increases, they represent an increasingly large percentage of the assemblage. Invertebrates designated as tolerant represent the 5-10% most tolerant taxa in a region. In a sense, they occupy the opposite end of the spectrum from intolerant taxa.

Percent predator. Predator taxa represent the peak of the food web and depend on a reliable source of other invertebrates that they can eat. Predators may have adaptations such as large eyes and long legs for hunting and catching other animals. The percentage of animals that are obligate predators provides a measure of the trophic complexity supported by a site. Less disturbed sites support a greater diversity of prey items and a variety of habitats in which to find them.

Percent dominance (3 taxa). As diversity declines, a few taxa dominate the assemblage. Opportunistic species that are less particular about where they live replace species that require special foods or particular types of physical habitat. Dominance is calculated by adding the number of individuals in the three most abundant taxa and dividing by the total number of individuals collected in the sample.